QUALITY SORTING AND NETWORKING: EVIDENCE FROM THE ADVERTISING AGENCY INDUSTRY

by

Mohammad Arzaghi *
Brown University

CES 05-16 October, 2005

Quality Sorting and Networking: Evidence from the Advertising Agency Industry*†

Mohammad Arzaghi‡
Brown University

First Draft: November 2004
This Draft: June 2005

Abstract

This paper provides a model of knowledge sharing and networking among single unit advertising agencies and investigates the implications of this model in the presence of heterogeneity in agencies' quality. In a stylized screening model, we show that, under a modest set of assumptions, the separation outcome is a Pareto-undominated Nash equilibrium. That is, high quality agencies locate themselves in a high wage and rent area to sift out low quality agencies and guarantee their network quality. We identify a necessary condition for the separating equilibrium to exist and to reject the pooling equilibrium even in the presence of agglomeration economies from networking. We derive the maximum profit of an agency and show the condition has a directly testable implication in the empirical specification of the agency's profit function. We use a sample of movers—existing agencies that relocate among urban areas—in order to extract a predetermined measure of their quality prior to relocation. We estimate the parameters of the profit function, using the Census confidential establishment-level data, and show that the necessary condition for separation is met and that there is strong separation and sorting on quality among agencies in their location decisions.

Keywords: Advertising; Agglomeration; Industrial Concentration; Business Services; Discrete Choice; Knowledge Spillovers; Learning; Location Decision; Poisson Regression; Nested Logit; Screening; Separating Equilibrium; Sorting
JEL classification: D82, D83, D85, L25, L84, M37, R12, R30

*I am very grateful to Vernon Henderson, Rajiv Vohra, and Andrew Foster for all of their advice and support. I also thank David Weil, James C. Davis, and others at Brown for their many insights and suggestions. All remaining errors are my own.

†The research in this paper was conducted while the author was Special Sworn Status research associate of the U.S. Census Bureau at the Boston Census Research Data Center (BRDC). Research results and conclusions expressed are those of the author and do not necessarily reflect the views of the Census Bureau. This paper has been screened to insure that no confidential data are revealed.

‡Department of Economics, Brown University, Box-B, Providence, RI 02912, USA; marzaghi@brown.edu.

1 Introduction

Advertising agencies are disproportionately concentrated in large and expensive cities, such as New York and Chicago. These large clusters comprise higher quality and bigger agencies than elsewhere. Why is such a pattern predominant in the clustering of advertising agencies? While one might suppose that access to media outlets, quality labor, and headquarters (advertisers) could explain this pattern, it appears that much of the clustering pattern is induced by market segmentation among agencies, where high quality agencies use costs as a device to sift out low quality agencies and thus guarantee their network quality.

Extensive research in economics has focused on the role of informal institutions and non-market interactions in shaping economic activities. In recent years, the literature on coalition formation[1] and network formation[2] has been particularly promising in order to rigorously analyze and formulate such non-market interactions among firms and individuals. This literature is especially significant, because it can reveal processes through which some of the most important externalities are generated, internalized, and dealt with by economic agents in the absence of a formal market structure (Coase 1960; Maskin 1994). For example, the coalition formation studies explain how knowledge sharing activity can be a self-enforcing strategy and can lead to research and development joint ventures among the high-tech firms even in the presence of pervasive free-riding problems (Yi and Shin 2000) or to informal and formal networking among advertising agencies despite severe competition among them (Arzaghi 2005). Thus, it must be no surprise that the study of learning and knowledge spillovers has been one of the most active areas of research in the above literature (Bala and Goyal 1998; Ellison 1993; Ellison and Fudenberg 1993; Yi 1997 to name a few). This specifically interests urban economists, because they, among others, have hypothesized learning and knowledge spillovers as the major sources of agglomeration economies, even though the micro-foundation behind such externalities has been mostly overlooked (see Duranton and Puga 2004 for a review). Jaffe, Trajtenberg, and Henderson (1993) provide convincing empirical evidence of such externalities and show that knowledge is bound to its location of origin. They conclude that knowledge spillovers quickly dissipate with distance, a point that is emphasized by Rosenthal and Strange (2001). Nevertheless, a critical ingredient of knowledge sharing activities has been ignored. In a world with heterogeneous agents, a contact with different types of agents provides different levels of benefits and may involve different costs of initiating and maintaining a

[1]See Bloch (1997) for a review of the literature, and Ray and Vohra (2001) and Yi and Shin (2000) for recent works. In this literature, the possibility of pre-game communications provides a host of attractive cooperative strategies for the players. A strand of the literature focuses on various refinements of Nash equilibrium by allowing group deviations, such as strong Nash or coalition proof Nash criteria. Others try to formulate the pre-game communication process into a larger and more complicated non-cooperative game and employ the sequential rationality concept to solve it.

[2]See Goyal and Moraga-Gonzalez (2002) for a review of literature. Wasserman and Faust (1994) provide an exhaustive summary of research on networks in the social sciences. For narrative evidence of existence and importance of networks, see Gladwell (2002). In economics, the discussion on networking in game theory goes back to Mayerson's (1977) research on graph theory. However, it is Jackson and Wolisnky (1996) who revived the networking idea and established the fundamental conflict between efficiency and stability of networks.

contact.

In this paper, we provide a model of knowledge sharing and networking[3] among small and creative businesses, specifically single unit advertising agencies,[4] and investigate the implications of this model in the presence of heterogeneity among the agencies. We not only establish the general scale effects from networking (localization effects), but also the possibility of sorting on quality. We show that, under a modest set of assumptions, higher quality agencies have an incentive to separate themselves from lower quality agencies. Given that the quality of an agency is private information, this results in the concentration of high quality agencies in locations with higher general costs (wage and office rent combinations). We show in a stylized screening model, with two types of agencies, that the separation is a Nash equilibrium given that a high enough cost differential exists. In so doing, we derive the maximum profit of an agency and show that the necessary conditions for the existence of a separating equilibrium have strong implications for the empirical specification of the agency's profit function. We estimate the parameter of the profit function using the Census confidential establishment level data. We show that the networking effect and separation are predominant factors in the profit function of agencies and, consequently, in the location decisions of new agencies and movers.[5]

Our initial intuition was that the profits of agencies may be heavily influenced by proximity to headquarters and media outlets (Silk and Brendt 1994; King, Silk, and Brendt 2003), because headquarters are the main buyers of advertising and mass media are the main suppliers of advertising space. The data, however, reject the idea that the proximity to headquarters is of any consequence for an agency's profits. It also shows that proximity to media is unimportant. Our estimations, however, highly support sorting on quality among agencies and networking as major determinants of agencies' profits, and, consequently, of their location decisions.

The next section explores the characteristics of the advertising agency business in the United States. First, we explain why the concepts of networking and knowledge sharing are particularly important in advertising. Then we provide some exploratory evidence of segmentation and sorting on quality among agencies. Section 3 describes the advertising agency's technology under the local networking hypothesis. Using a stylized screening model, we show that the separation is the likely outcome and the necessary

[3]Note that in the rest of this paper, we will refer to a coalition of agencies as a network. In fact, a group of independent agencies under a holding company is called a network in the industry. This is different from the terminology that is adapted in the theoretical literature on "network", which defines "network" as the total structure of links among the players. In a sense, we impose a specific condition on the structure of a network. That is, all the members within a coalition are connected (complete graph within a group) and there is no link to the members of other coalitions.

[4]See Arzaghi (2005) for the rationale of using advertising agencies and the micro-foundation model of knowledge sharing among advertising agencies. The paper provides a model of network formation among advertising agencies. It shows that forming networks and sharing knowledge within a network is a strategic response of agencies to a search and matching problem in the market for advertisement.

[5]An agency is called a "mover" if it is an existing advertising agency that has relocated across urban areas.

condition for separation has a testable implication in our empirical specifications. In section 4, we derive the empirical model based on the random profit formulation and show that it lends itself to four different empirical specifications as we entertain different assumptions about the structure of error terms and remedy the potential identification problems. In section 5, we briefly introduce and explain our data and the list of socioeconomic, historical, and topographical instruments that are used in estimation. Section 6 summarizes the results and explains their relevance, and section 7 offers our conclusions.

2 Advertising Agency Business in the U.S.

Historically, advertising outputs make up 2 to 3 percent of the U.S. economy. In 1997, advertising expenditures totaled \$187 billion, about 2.2 percent of GDP (NIPA, BEA). The 1997 Economic Census reports that \$96 billion, more than half of advertising expenditures, are handled by advertising agencies. This includes \$71 billion in media billings, \$18 billion in billings for outsourced materials and services, and \$7 billion in fees and services. Of these, only \$17 billion involves advertising agency receipts.[6] In short, while the contribution of advertising agencies to the national income is not large, they manage and shape a sizeable portion of the U.S. economy.

Advertising agencies perform two key roles. The historical role, which is still important, is as intermediaries or matchmakers between advertisers and the media.[7] They used to receive a 15 percent commission on media billings for their services.[8] Over time, that role led to the emergence of a second function for which agencies are better known today—to design and produce advertising campaigns. This puts agencies at the creative front of the advertising business. In this respect, the advertising agency business is formed to digest the advertiser's marketing problems and to produce ideas and creative campaigns that increase the advertiser's sales, profits, and/or reputations.[9] Usually, agencies operate with internal teams that are a collection of creative, accounting, and research personnel. A team typically provides a full range of services for client-marketing research, creation of an ad campaign, production of an ad, media negotiation and placement, and promotional suggestions. All of these activities are highly labor intensive, especially developing

[6] This includes \$7.8 billion in commissions on media billings, \$2.2 billion in commissions on billings for outsourced materials and services, and, of course, \$7 billion in fees. This puts the media commission rate at 11 percent in 1997.

[7] An advertiser is a firm that demands advertising campaign in order to market its products, cast an image, form reputation, and so on.

[8] Today the average rate is closer to 10 percent.

[9] There are issues regarding the effectiveness and rationale of advertising. It is a separate topic that has drawn both theoretical and empirical attention in the field of Industrial Organization. A survey of that literature demands a separate paper. Good but short discussions of the theory are provided in Shy (1995) and Tirole (1988). Bagwell (2003) provides an exhaustive survey of the economic analysis of advertising, and Bagwell (2001) includes the comprehensive collection of empirical and theoretical studies on advertising in economic literature. For a good and clean, although outdated, survey and discussion of theory and empirical results, we refer you to Schmalensee (1972). Here, we brush those questions aside by assuming that advertising is needed by advertisers (manufacturing and other industries) as a part of their input to production of final goods. We do not attempt to justify the need for advertising or its effectiveness.

new and creative ideas for ads or promotional suggestions. About 72 percent of agencies' costs are labor costs (Economic Census, 1997). The only other tangible cost is the rent and lease of office spaces.

The advertising agency business is built on information. Information about the advertiser's product, the potential competitors, the kind of people who consume the product and their characteristics, the best possible medium through which to reach them, local market information, experience with all type of media, where to find the best copywriter and art director, and what are the latest advancements in the internet technology (interactive advertisement) help an agency to create a successful proposal that acquires an account. This is perhaps the main reason that most of the larger agencies have a research department, and purchasing information from outside venders is a common practice among all sizes of agencies. However, not all kinds of information are easily available to be purchased. It is true that the characteristics of products or the market are often available for the right price, but detailed information such as the effectiveness of a certain line of advertising for a category of products, who is the best copywriter or art director for the job, the personality of an advertising manager of a client company, or tips that a certain advertiser is seeking a new agency to handle its account is not. More importantly, there is no market for new, creative, and untried ideas. Thus, similar to research and development (R&D) in Jaffe et. al (1993), the flow of information and sharing knowledge and creative ideas are crucial ingredients for a successful business in the advertising industry.

The knowledge sharing happens within and across agencies through formal and informal networks. There is substantial evidence of formal networking in the industry. Hameroff (1998) lists eleven large national and international networks, and *Advertising Age* in its "2003 Agency Report" identifies 21 independent agency networks. Members pay dues and attend network meetings several times a year. These networks have full-time managers, and members are typically located in different cities or even countries. The members exchange information on all aspects of business from accounting practices, fees, and commission rates to the state of the art developments in animation pictures and internet technology. Other arrangements that resemble formal networks in the advertising business are "holding companies." A holding company (sometimes called a network) is a collection of functionally independent agencies (in H-form corporate) that can provide the support of a formal network while it abates the problem of account conflict, something that a single agency could not do. The fact that most holding companies have multiple offices in the same geographical locations is hard to explain in the absence of the networking phenomenon.[10] However, one should be cautious with this interpretation of holding companies, because a holding company may provide its members with services beyond networking support, such as a financial cushion, flexibility in resource allocation, better bargaining power in media buying, access to the capital market, and wider geographical reach; yet, the members are burdened with the additional costs of negotiating and enforcing the formal contracts. These make it very

[10] Multi-unit agencies in Manhattan have about 3 establishments on average (Arzaghi and Henderson 2005).

difficult, if not impossible, to solely identify the networking and knowledge sharing benefits (and possible sorting) among the members of formal networks (or holding companies).[11] In addition, formal networks include a small fraction of all advertising agencies, though admittedly they include the larger ones. Even though these formal networks are interesting, we focus on single unit firms in this paper to avoid the complexity of the multi-unit firms.

While formal networks need costly managers and other organizational staff (or legal arrangements in the case of holding companies) and the members are usually geographically disperse, the informal networks among smaller agencies (or their employees) are usually based on everyday personal contacts and formed within a close proximity.[12] These informal networks are the focus of this paper. In the advertising agency business, informal networking plays a big role because the industry is dominated by single unit and small establishments. Single unit firms (SUs) account for 90 percent of establishments and 55 percent of sales in the industry, compared to 81 percent and 45 percent for all other business services (Economic Census, 1992). Furthermore, multi-unit firms with more than 10 establishments only account for 22 percent of sales, in contrast to 33 percent for all other business services. An average single unit advertising agency (5.9 employees) is about half the size of SU firms in other business services (11.4 employees).[13] In short, the advertising agency business is dominated by small single unit agencies and informal networking is crucial to these small agencies.

Networking appears to occur at the design stage of advertising campaigns when an agency receives a "request for proposal" (RFP) for a new advertising campaign (Arzaghi 2005). The agency draws upon its network (who, given the odds, are unlikely to have received the same RFP) to put together the best response to the request. This improves the quality of the proposal and increases the relative chance of winning the account. This highlights the advantages of belonging to a network with high quality members. In our model in the next section, we intend to capture these effects by assuming that the productivity of an agency is not only a function of its own quality but also a function of its networking decisions and the quality of its network's members.

On the demand side for advertising, the headquarters are the main buyers of advertising products,[14] with

[11]Each holding company has a unique legal structure that defines the boundaries and responsibilities of the agencies. In the same way, every formal network has specific themes and goals.

[12]Interviews with dozen executives in different agencies in New York and Chicago support the existence and prevalence of networks and professional friendships in the advertising agency industry. They express that these contact are extremely important to their business. In addition interviews show that in about 90 percent of cases, the last contact was located within the same city and about half of them were in a walking distance.

[13]Source: The 1992 Economic Census, Subject Series, Establishment and Firm Size, Census Bureau (1995). By "business services", we are referring to SIC-73.

[14]The break down of advertising demand among advertisers is: automobile industry 18 percent, retail, department and discount stores 15 percent, and then another 36 percent divided evenly among the eight following industries: movies, cosmetics and toiletries, drugs, food, financial services, restaurants, airline and travelling, and telecommunications (Taylor Nelson Sofres, Consulting Media Research). The share of national advertisers is about 55 of total advertising .

advertising expenditures of headquarters as high as 37 percent of headquarters wage bills, compared to 13 and 15 percent on accounting and legal expenditures, respectively (Henderson 2004). Nevertheless, there is no evidence that agencies are significantly affected by the geographical location of headquarters. As Davis and Henderson (2004) point out, geographically, headquarters are distributed evenly, and most of headquarters are located close to their production plants, as opposed to agencies, which are highly concentrated in a few urban areas. The fact that, for larger accounts, agencies typically dispatch a team to work with advertisers and multi-unit agencies could afford to have regional offices close to advertisers[15] (Hameroff 1998; Clark 1988; McDonough 2003), could explain the lack of correlation. In a sense, advertising agencies are in the business of producing fresh ideas in the most vibrant advertising locations and exporting advertising services (Arzaghi and Henderson 2005).

There is a strong relationship between advertising agencies and media. In 1997, about half of agencies' incomes were from the commissions on media billings (Table 1). However, on average, SUs are less dependent on the media. Only 28 percent of SUs' incomes were from commissions on media billings compared to 48 percent for multi-unit agencies (Table 1). Media commissions are defined as negotiating contract fees and/or commissions for placing advertising in radio, television, newspapers, periodicals, and other media. It is clear from this definition that these tasks involve business and marketing skills and require extensive and up-to-date knowledge of broadcasting and publishing businesses, but it virtually involves no creative activity (Wells et al. 2003). Media placement became so specialized in the late 70s and early 80s that media buying became a separate line of business in order to provide services to other advertising agencies. Almost all MUs have a separate (internal) department that deals with media buying activities (McDonough 2003; Wells et al. 2003; Hameroff 1998; Nicosia 1971). As opposed to multi-unit agencies, single unit agencies rely on incomes from providing advertising materials and services, such as graphic design, photography, artwork, plates, printed materials, and market research. These may be supplied and performed by the agency or the agency may only play the role of a broker for these services and receive commissions. By definition, these activities are more creative in nature. In 1997, on average 25 percent of SUs' incomes were from these creative activities compared to only 10 percent for MUs. This emphasizes the roles of large and media-intensive accounts in MU agencies' business in contrast to the creative and service-oriented tasks of SU agencies.

The geographical distribution of advertising agencies in the U.S. is highly uneven. Table 1 shows the basic characteristics of the advertising agency business in the U.S. and the top three PMSAs, ranked by their total receipts. New York accounts for 24 percent of national advertising agency income, 20 percent of employment, and 29 percent of all media billing in the U.S. All these measures show that New York is

[15] For example, McCann-Erickson has offices in Detroit and San Francisco to handle its major accounts with GM and Microsoft, respectively.

about three times larger than the second largest advertising center: Chicago.[16] Also, more than a third of all advertising employment is concentrated in the three largest centers. These advertising centers carry out a disproportionate share of the industry. The share of New York PMSA in advertising employment is much higher than New York's share in national employment (2.7 percent) and also much higher than New York's share of employment in professional and scientific services (7.5 percent). Moreover, most advertising activities are located in the central business districts (henceforth CBDs) of these cities. For example, in 1992 more than 97 percent of all advertising employment in New York PMSA was located in Manhattan.

[**insert Table 1**]

Figure 1 shows that advertising agencies are disproportionately located in counties with larger employment. The shares of advertising establishments among all establishments, measured by location quotients,[17] are strongly related to the scale of economic activities. The larger economic activities usually corresponds to higher office rents and wages, and, as we will see later, agencies may use these higher costs as a device to screen for the quality of the networking possibilities. We explore this idea in the following paragraphs.

[**insert Figure 1**]

Single unit agencies in New York in general are larger and handle more business than elsewhere. Table 1 shows that the average sales/receipts of SUs in New York is 2.2 times higher than the national average. The ratio is 1.8 for the average employment of single unit agencies. In addition, their activity per employee is about 20 percent more than the national average, and their salaries are about 60 percent above the national average.[18] We can observe the same pattern for other advertising centers. For example, the average size of a single unit agency in Chicago (measured by employment) is about 12 percent and in Los Angles about 20 percent above the national average. A simple statistical test clearly supports that there are significant size differences among SU agencies across the U.S. metro areas.[19] In the advertising business, the most important sign of success is growth (Hameroff 1998). The idea is that an agency must capitalize on its ability and ideas

[16]This is based on the total for both MU and SU agencies using the 1997 Economic Census. The total sales of SU agencies in the New York CBD is two times larger than the second largest advertising center, however the second largest advertising center based on SUs' sales is Los Angles and not Chicago (Table 1).

[17]The location quotient of the advertising agency business is defined for CBD k as the following ratio

$$LQ_k = \frac{\frac{t_{ik}}{t_k}}{\frac{t_i}{t}}$$

where, t_{ik} is the number of advertising establishments at location k, t_i is the number of advertising establishments nationally, t_k is the number of all establishments at location k, and t is the number of all establishments nationally.

This provides a more appropriate measure of concentration than simple shares in the national advertising business, since it measures how disproportionately advertising agency establishments are concentrated at every location as compared to overall economic activities in the location.

[18]Of course, the higher wages partially reflect the general increase in living costs due to urban congestion (the traditional story of urban rent gradients, Henderson 1988). Nevertheless, the higher wages also indicate that the average employee and the average agency are more productive in New York than in other places. This may be due to the selection of more productive agencies and personnel into New York City.

[19]We use ANOVA test on the size of SU agencies across the PMSAs. It decisively reject the null hypothesis ($F(302, 8630) = 1.95$ and P-value= 0.0000)that the size of SU agencies across metro areas are random draws from the same distribution.

to acquire as many accounts as it can. In this sense, the size of an agency is a measure of its quality.[20] Figure 2 indicates that larger agencies are located in CBDs that have more agencies. This resembles the same phenomenon that usually appears in the sorting literatures. Higher quality agencies normally belong to larger communities.

In Figure 3, we show the average agency size against office rent for all 106 CBDs with available office rent for square foot of class "A" office spaces. This figure clearly shows the segmentation of higher quality agencies in locations with higher rents. Because agencies comprise a negligible portion of demand for office spaces, we think of rents as a given (exogenous) price of office space to advertising agencies. This is a clear indication that higher quality agencies use costs (rents and wages) to separate themselves from lower quality agencies.

[insert Figure 2 and 3]

An average single unit agency stays in the market for about 6 years. The birth rate is above 50 percent over a five year period in the industry. This means that in less than 15 years more than 90 percent of agencies are replaced with new faces. This high turnover causes constant changes in the composition of the networks. This highlights the importance of separation (or locating in high costs areas in order to sift out low quality agencies) as the (only) way to effectively screen the quality of networking in the highly uncertain world of advertising agencies.

The data shows that the geographical distribution of SU agencies has changed since 1977.[21] We can statistically reject that even the geographical distribution of SU agencies in 1987 is the same as that in 1997. Clearly, the differences are caused by shifts in the pattern of births during 77-97. Given the fact that few agencies have survived since 1977 and that the geographical spread of SU agencies has changed (interestingly, this is not true for MU agencies), we are confident that the births in the 1992-97 period (or even 1987-97) are not mainly driven by the constant and long-lasting attributes of locations. However, remnant agencies provide enough correlation between historical and current attributes of local SU agencies. Hence, we feel that historical attributes of SU advertising agencies could provide good instruments for their current counterparts.

In the next section, we present a model of local interactions and show that under a modest set of assumptions the separating equilibrium exists, where high quality agencies use costs as a device to sift out

[20] More broadly, Jovanovic 1982 shows a positive relationship between the quality and the size of an establishment.

[21] We use the contingency table approach and Pearson χ^2 tests for the homogeneity of the geographical distributions. This is implemented using a Generalized Linear Model (GLM). The simple explanation of this method is that it regresses the log of the number of events in each cell on the log of marginal events associated with that cell. In other words, if the distributions are similar (homogeneous) then we have $p_{ij} = p_i p_j$ or $\frac{n_{ij}}{N} = \frac{n_i}{N}\frac{n_j}{N}$. This leads to a log linear formulation of

$$\ln n_{ij} = \ln n_i + \ln n_j - \ln N$$

and the sum of the squared residuals from the above regression provides us with Pearson statistics.

low quality agencies and thus guarantee their network quality.

3 A Model of Local Interactions

We consider networking among heterogenous agencies given that the quality of an agency is private information. We assume that there are a finite, but large, number of available locations (i.e., cities) from which agencies can choose. These locations provide a wide spectrum of wages and office rents. Establishing contacts (networking) enhances the productivity of agencies by providing new perspectives and fresh ideas and by exchanging expertise and knowledge (Arzaghi 2005); however, each contact exhausts a fraction of the intellectual ability of the agency in the form of efforts to initiate the contact, to learn about the other agency, and to maintain the contact. A contact is a reciprocal relationship. This means that both sides are affected by the contact. The benefits from a contact depend on the quality of both participants and the agency's effort. The marginal cost of effort increases with an agency's quality. We assume that contacts are only possible among agencies at the same geographical location, such that the choice of location determines the set of potential contacts.

Given that there is heterogeneity among agencies in their quality and their networking costs, we expect that the agencies' actions cause separation and sorting based on quality. In this respect, we follow the traditional screening models (Rothschild and Stiglitz 1976; Riley 1979a,b). We show that high quality agencies have an incentive to separate from lower quality agencies and can achieve that by choosing locations with high wages and office rents. This model differs from the traditional screening models in two respects. First, as opposed to the traditional models, there is no distinct principal or agent in the model: every agency plays the role of a principal and an agent. Second, there are agglomeration economies from networking among agencies that further complicate the matter.

In the following, we first show that the profit of an agency is a function of the size and average quality of its network. Then using a stylized screening model with only two types of agencies, high and low quality, we study separating and pooling equilibria. We show that, under a modest set of assumptions, the separating equilibrium is a unique Pareto-undominated Nash equilibrium. A necessary condition to reject pooling equilibria and to support the separating equilibrium provides a testable implication for the parameters of the agency's profit function. We derive this implication in section 4 and test it in section 6.

3.1 General Setup and Assumptions

Consider an economy that comprises K distinct locations. After locating at location $k \in K$, an agency hires skilled labor and uses office space in the production of advertisements. That is, the agency i's production function at location k is

$$Y_{ik} = A_{ik} l_{ik}^{\alpha} s_{ik}^{\beta},$$

where A_{ik} is the productivity of agency i at location k, which includes the contribution of the agency's network. The wage and office rent at location k are w_k and r_k, respectively, and we assume that locations can be ranked based on the (w_k, r_k) pairs.[22] An agency's quality is identified by the quality of the entrepreneur or entrepreneurs who run(s) the agency's creative, account management, and other core activities. We indicate the quality of agency i with $q_i \in [0, 1]$. The quality of an agency is private information.

An agency can only establish contacts with other agencies at the same location.[23] We represent the value or quality of the contact between agency i with agency j at the location k with $q_{ijk} \in [0, 1]$. This is an increasing function of the quality of agency i and agency j. We assume that the contacts are reciprocal. That is, $q_{ijk} = q_{jik}$.[24]

Agency i's primary productivity level only depends on its own quality

$$A_{ii} = \bar{A} \rho q_i,$$

where ρ is a positive constant and $\bar{A}$ captures the general average Hick neutral productivity factor.[25] In other words, A_{ii} is the productivity of agency i in a vacuum. Given that $\rho > 0$, this assumption guarantees that agency i is productive even in the absence of networking. Agency i's productivity increases by the amount of $A_{ijk} = \bar{A} e_{ijk}^{\eta} q_{ijk}$ when it establishes a contact with agency j and allocates the effort of e_{ijk} to the contact. We assume that there is diminishing returns to effort, that is, $0 < \eta < 1$.[26] There are non-labor entrepreneurial (opportunity) costs to agency i of $\bar{A} c(e_{ijk})$ per unit of output associated with each contact.[27] Furthermore, we assume that

$$c(e_{ijk}) = c_i e_{ijk} + F,$$

where F is the fixed cost of learning the quality of the other agency and e_{ijk} is the level of (creative and/or intellectual) effort that agency i devotes to the contact only after the quality of the contact is revealed. Note

[22] The overall contribution of the wage and office rent to the profit function of an agency at location k can be captured by $\alpha \ln w_k + \beta \ln r_k$, where α and β are the labor and office space elasticities in the production function, respectively. This value provides the basis for the ranking of locations on their cost disadvantages.

[23] The required effort to initiate and maintain the contact between distant locations is prohibitively high. We specifically assume that contacts are limited within a metropolitan area.

[24] Specifically, we assume $q_{ijk} = (q_i q_j)^{1-\eta}$ in the next subsection. This function satisfies the above assumptions.

[25] In section 4, we introduce a random element as a part of the productivity of an individual agency to capture the idiosyncratic matching parameter between agency i and location k. In that regard, $\bar{A}$ is an average over this idiosyncratic element.

[26] An alternative approach is to assume that effort, e_{ijk}, affects the probability of establishing the contact in the following way

$$P_{ijk} = \frac{e_{ijk}}{1 + e_{ijk}}.$$

Therefore, the expected productivity gained is $\bar{A} \frac{e_{ijk}}{1+e_{ijk}} q_{ijk}$. This is a concave function of e_{ijk} and shows diminishing returns on effort.

[27] These are the opportunity costs of the time and effort of principal people, such as account managers or creative people, in the agency.

that other agencies are indistinguishable to an agency before it incurs the fixed costs of initiating contact and learning about them. The marginal cost of the effort, c_i, is positive and varies among the agencies, with a presumption that higher quality agencies have a higher marginal cost per unit of their entrepreneurial efforts.

Thus, the overall contribution of the contact with agency j at location k to the productivity of agency i is

$$\begin{aligned} A_{ijk} &= \bar{A}\left(e_{ijk}^{\eta} q_{ijk} - c\left(e_{ijk}\right)\right) \\ &= \bar{A}\left(e_{ijk}^{\eta} q_{ijk} - c_i e_{ijk} - F\right). \end{aligned}$$

3.2 Agency's Production Schedule

The ex-post profit of agency i, if it chooses location k, is

$$\Pi_{ik} = A_{ik} l_{ik}^{\alpha} s_{ik}^{\beta} - w_k l_{ik} - r_k s_{ik},$$

where l_{ik} and s_{ik} are labor and office space hired by the firm i at location k, respectively, and

$$A_{ik} = \bar{A}\left[\rho q_i + \sum_{j}^{N_k}\left(e_{ijk}^{\eta} q_{ijk} - c_i e_{ijk} - F\right)\right]$$

is the productivity of agency i in location k.

When agency i chooses a location, it incurs the rent and wage at the location, and in return it is able to contact agencies in the area. The agency learns the value of contacts with agency j at the location k, q_{ijk}, at the fixed cost of F per unit of output. Given that cost F is sunk, the agency allocates an optimal and non-negative level of effort e_{ijk} to maintain the contact with agency j.

We can write the profit function of agency i at location k as

$$\Pi_{ik} = \bar{A}\left[\rho q_i + \sum_{j}^{N_k}\left(e_{ijk}^{\eta} q_{ijk} - c_i e_{ijk} - F\right)\right] l_{ik}^{\alpha} s_{ik}^{\beta} - w_k l_{ik} - r_k s_{ik},$$

and the optimal level of effort for contact with agency j at location k is given by

$$e_{ijk}^{*} = \left(\frac{\eta}{c_i}\right)^{\frac{1}{1-\eta}} q_{ijk}^{\frac{1}{1-\eta}}. \tag{1}$$

By substituting the optimal level of effort, we can show that the contribution of a contact with agency j is

$$A_{ijk}^{*} = \bar{A}\left((1-\eta)\left(\frac{\eta}{c_i}\right)^{\frac{\eta}{1-\eta}} q_{ijk}^{\frac{1}{1-\eta}} - F\right).$$

For the purposes of simplicity, we assume that $q_{ijk} = (q_i q_j)^{1-\eta}$. This satisfies our main assumption that the quality of a contact is an increasing function of the quality of both participants, and also that it belongs to the interval $[0,1]$. Thus the ex-post profit function at location k is

$$\Pi_{ik} = \bar{A}\left[\rho q_i + \sum_{j}^{N_k}\left((1-\eta)\left(\frac{\eta}{c_i}\right)^{\frac{\eta}{1-\eta}} q_i q_j - F\right)\right] l_{ik}^{\alpha} s_{ik}^{\beta} - w_k l_{ik} - r_k s_{ik}.$$

The ex-post maximum profit at location k is

$$\Pi_{ik}^{*} = (1-\alpha-\beta)\, A_0 N_k^{\frac{1}{1-\alpha-\beta}} \left[\frac{\rho q_i}{N_k} + (1-\eta)\left(\frac{\eta}{c_i}\right)^{\frac{\eta}{1-\eta}} q_i \bar{q}_k - F\right]^{\frac{1}{1-\alpha-\beta}} w_k^{-\frac{\alpha}{1-\alpha-\beta}} r_k^{-\frac{\beta}{1-\alpha-\beta}}, \tag{2}$$

where $\bar{q}_k = \frac{1}{N_k}\sum_{j}^{N_k} q_j$ is the average quality of agencies at location k and $A_0 = \left(\bar{A}\alpha^{\alpha}\beta^{\beta}\right)^{\frac{1}{1-\alpha-\beta}}$ is a constant.

The optimal level of employment is given by

$$w_k l_{ik}^{*} = A_0 N_k^{\frac{1}{1-\alpha-\beta}} \left[\frac{\rho q_i}{N_k} + (1-\eta)\left(\frac{\eta}{c_i}\right)^{\frac{\eta}{1-\eta}} q_i \bar{q}_k - F\right]^{\frac{1}{1-\alpha-\beta}} w_k^{-\frac{\alpha}{1-\alpha-\beta}} r_k^{-\frac{\beta}{1-\alpha-\beta}}. \tag{3}$$

3.3 Location Decisions and Separation

We have specified how the profit of an agency is affected by the quality of its neighbors. Clearly, this factor determines the location decisions of agencies in the first place. In the following section, we motivate the idea of the separation using a stylized model, where there are only two types of agencies and there are a large number of locations that provide agencies with a wide spectrum of wage and office rent combinations. Clearly, this microcosm is intended to motivate our empirical specifications and to help us interpret the results.

Quality of Agencies—Until now, we have assumed that agencies are different in quality, but we were not specific about their classification. We capture the agency's quality, however, with two parameters in the production function. First, q_i captures the quality of the agency's contribution to a contact, that is, the effectiveness of the agency in networking. Second, we consider the non-labor marginal cost of effort per unit of output c_i (entrepreneurial opportunity costs per unit of effort) to maintain the contact. We assume that the marginal cost of the effort is higher for the agency with the higher quality. In other words, a unit of effort by high quality agencies is more productive and thus its opportunity cost is higher than that of a low quality agency.

Of course, there is a continuous spectrum of quality among the agencies. As we have assumed, q_i could take any value belonging to the $[0,1]$ interval, with the caveat that a higher q_i corresponds to a higher c_i. It is clear that an agency, regardless of its own quality, benefits more from networking with higher quality agencies. Given that the entrepreneurial effort is more costly to high quality agencies, they would rather better allocate their efforts for contacts with other high quality agencies. However, they cannot identify the quality of other agencies prior to contact and before spending the fixed cost F. Under these circumstances, higher quality agencies have an incentive to sift out those of lower quality. In general, even ignoring the scale externalities in our model, this is a very complex problem with the possibility of multiple equilibria, especially when the prices (wages and rents) are treated as endogenous to the sorting process (Epple and Sieg 1999; Berry, Levinsohn and Pakes 1995, 2004; Bayer, McMillan, and Rueben 2004). However, given that the agencies only account for very small fractions of the demand in the local labor and office space markets,[28] we think that wages and rents are not influenced by the location decisions of the agencies. Thus, for all practical purposes, wages and office rents are exogenous to the decisions of agencies,[29] and we treat agencies as price takers, where the rents and wages are given and vary across locations. Even though this simplifies the matter, the possibility of coordination among agencies and the networking scale effect on profits make the problem intractable for the continuum of quality.

Therefore, in our stylized model, we assume that there are only two types of agencies in the economy: high quality agencies or H types with q_H and c_H attributes, and low quality agencies or L types with q_L and c_L attributes, where $q_H > q_L$ and $c_H > c_L$. The first inequality implies that contacts with the high quality agencies are more productive. In the second inequality, we assume that the marginal cost of the effort by the high quality type is higher than that of the low quality type.

Game—We think that, in the real world, there is no restriction on the location decisions of agencies and they can easily relocate. Thus, we employ rules that impose the least amount of restrictions on the location decisions of agencies. The game that we have in mind is an open membership game (or free mobility game) as introduced in Yi (1997) and used in a networking context by Arzaghi (2005).

In this static game, each of M agencies simultaneously announces an "address" among K possible addresses. Let $\{a_1, ..., a_K\}$ be a set of K addresses, so that player i's strategy, σ_i, is one of the $a_1, ..., a_K$ addresses. In this specific case, agencies that pick the same address are in the same physical location (cities). The resulting numerical structure is $\mathbf{N} = \{N_1, ..., N_K\}$, where N_l is the number of agencies at location l. The agencies at each location establish contacts with all other agencies in the same location. Quality of agencies are revealed at this stage.

Nash Equilibria—It is clear that the numerical structure $\mathbf{N} = \{N_1, ..., N_K\}$ presents a Nash equilibrium as long as no agency at any location finds it profitable to deviate in order to stand alone or to join agencies in another location.

[28] Even the large and disproportionate concentration of agencies in Manhattan, New York, only accounts for less than 0.7 percent of all establishments in Manhattan.

[29] Even though wages and rents (prices) are not endogenous in the traditional sense, where wages and rents are influenced by the location decisions of economic agents (here agencies), there could be several reasons to suspect that profits of agencies and wages and rents are influenced by unobserved underlying factors. We will attend to this issue in detail in the empirical section of this paper.

We limit our attention to pure Nash equilibria, both pooling and separating. A separating equilibrium, in this context, is one in which high quality agencies choose the same address and collocate at a high wage and rent area to separate themselves from low quality agencies. In a pooling equilibrium, all agencies choose a single location.

Networking Benefits.—We assume that a contact with one's own quality type enhances productivity. That is,

$$(1-\eta)\left(\frac{\eta}{c_H}\right)^{\frac{\eta}{1-\eta}} q_H^2 \quad > \quad F \tag{4}$$

$$(1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_L^2 \quad > \quad F. \tag{5}$$

These conditions simply say that H types at the same location benefit from networking with one another, and L types benefit from networking with other L types. In fact, given these conditions, there can be only one occupied location (cluster) of each type. To show this, let us assume that L types are divided between two locations (one or both locations can include some H types). In general, an L type in one location has a higher profit (we break the tie by assuming they prefer the lower cost area in the case of a tie). A deviation by a single L type from the location with lower profit to the location with higher profit always increases profit. The profit increases because adding an L type to the cluster always increases the productivity and, hence, the profit of all the L types in that location. The same argument can be used for H types. We summarize this in the following proposition.

Proposition 1 *In any Nash equilibrium, each type occupies one location at most.*

Thus, only two types of Nash equilibria are possible: separating and pooling. Note that all pooling outcomes can be Pareto-ranked, where the Pareto-undominated outcome is one in which all agencies collocate at the lowest cost location. The separating outcomes also can be Pareto-ranked. We discuss this matter later in the section as we investigate the separating equilibria.

Given (5), it is clear that L types benefit from contacts with H types, because

$$(1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_H q_L > (1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_L^2 > F \tag{6}$$

as long as $q_H > q_L$.

A necessary condition for separation is that H types do not benefit from contacts with L types. That is,

$$(1-\eta)\left(\frac{\eta}{c_H}\right)^{\frac{\eta}{1-\eta}} q_H q_L < F. \tag{7}$$

In other words, the benefit of establishing a contact with a low type does not offset the fixed and variable costs of effort by a high quality agency. Otherwise, the pooling equilibrium in which all agencies collocate at the lowest cost area Pareto-dominates any separating equilibria. That is, any cluster of H types and L types achieve lower productivity than if they collocate at the lowest cost area.

Proposition 2 *No separating equilibrium can be a Pareto-undominated Nash equilibrium if (7) is not satisfied.*

Using conditions (5) and (7), one can see that

$$\frac{q_H}{q_L} < \left(\frac{c_H}{c_L}\right)^{\frac{\eta}{1-\eta}}$$

or

$$\frac{\Delta \ln c}{\Delta \ln q} \equiv \frac{\ln c_H - \ln c_L}{\ln q_H - \ln q_L} > \frac{1-\eta}{\eta}. \tag{8}$$

This relationship between the unit costs of effort and the quality of the two types of agencies is a necessary condition for the existence of a separating equilibrium. In other words, for any value of $\frac{\Delta \ln c}{\Delta \ln q} < \frac{1-\eta}{\eta}$, the high types benefit from networking with low types and the pooling equilibrium prevails. This necessary condition for the existence of a separating equilibrium has a directly testable implication in our empirical specification, which we will discuss in section 4.

Note that the costs of initiating the contact and learning about the quality of another agency, F, is sunk at the production stage. Thus, an H type agency allocates an effort e^*_{ijk} to the contact with an L type as long as $A_{ik} > 0$, even though it reduces its overall productivity. Clearly, given (7), no H type agency would choose a location where all contacts are of low quality.[30] Also, there is no incentive for a high quality agency to initiate a contact if it knows in advance that the other participant is of the low type. This means that, in general, high quality agencies have an incentive to stay away from low quality types. On the other hand, L types have a great incentive to be close to high quality agencies and to establish contacts with them; that is why a separating outcome is expected in the model.

Separating Equilibria—To establish that a separating outcome is a Nash equilibrium, we need to demonstrate that the traditional conditions of *individual rationality* and *incentive compatibility* constraints are satisfied for each type.

Given that an agency's quality is private information, *ex-ante*, the only way for high quality agencies to sift out those of low quality is to choose a location with a combination of wage and office rent that is not affordable (profitable) for low quality agencies, which leads to the incentive compatibility constraint of L types. That is, an L type should not benefit form deviating and locating with H types in a high wage-rent location. However, if the wage-rent combination is too costly, then an H type prefers to stand alone in a

[30]Given our assumptions, the agency can always choose to stand alone at a location with a wage-rent combination almost equal to the wage-rent combination at its current location. In this case, the agency's productivity is not affected negatively by low quality contacts. Here, we use the assumption that there are a large number of locations that provide a spectrum of wage-rent combinations.

lower cost location rather than locating in the high wage rent location with other H types. Put another way, individual rationality constraint of H type must hold. The other two constraints—the individual rationality of L types and the incentive compatibility of H types—are satisfied if the above two constraints hold.

Without a loss of generality, let us label the lowest wage-rent combination as (r_L, w_L)[31]. Then, a separating equilibrium exists if there is a wage-rent combination (r, w) such that

$$\frac{\rho q_L + N_H\left((1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_H q_L - F\right)}{\rho q_L + N_L\left((1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_L^2 - F\right)} \leqslant \left(\frac{w}{w_L}\right)^{\alpha}\left(\frac{r}{r_L}\right)^{\beta} \tag{9}$$

and

$$\frac{\rho q_H + N_H\left((1-\eta)\left(\frac{\eta}{c_H}\right)^{\frac{\eta}{1-\eta}} q_H^2 - F\right)}{\rho q_H} \geqslant \left(\frac{w}{w_L}\right)^{\alpha}\left(\frac{r}{r_L}\right)^{\beta}. \tag{10}$$

A separating equilibrium consists of the set of strategies where all low types choose to locate at (w_L, r_L) and high types choose a location with (w, r) that satisfies the above constraints.

Constraints (9) and (10) are the incentive compatibility of the L type and the individual rationality of the H type, respectively. Given (5), the individual rationality constraint of the L type is always satisfied. That is because an L type loses valuable contacts if it deviates form choosing (w_L, r_L) and stands alone. Moreover, condition (4) and constraint (10) guarantee that an H type never prefers to choose (w_L, r_L) and collocate with all the L types. Thus, the incentive compatibility constraint of the H type also holds.

We can combine constraints (9) and (10) into a single constraint on the parameter of the model

$$\frac{N_H\left((1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_H q_L - F\right) - N_L\left((1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_L^2 - F\right)}{\rho q_L + N_L\left((1-\eta)\left(\frac{\eta}{c_L}\right)^{\frac{\eta}{1-\eta}} q_L^2 - F\right)} \leqslant \frac{N_H\left((1-\eta)\left(\frac{\eta}{c_H}\right)^{\frac{\eta}{1-\eta}} q_H^2 - F\right)}{\rho q_H}, \tag{11}$$

which guarantees the existence of separating equilibria. It is clear that if N_L is large, then constraint (11) holds regardless of other parameters of the model. Note that higher N_L makes deviation by an L type to join the H types less likely. It is also clear that as long as the benefit of networking among all the H types is larger enough than standing alone, the constraint (11) holds. Our goal here is just to establish that the constraint (11) can hold and, thus, a separating equilibrium exist, and not to investigate the full range of parameters that guarantees constraint (11).

We label the wage-rent combination that makes (9) an equality as (w_H, r_H). This is clearly the lowest wage-rent combination that H types can choose to effectively sift out L types. Of course, (w_H, r_H) is the

[31]This means the wage-rent combination (w_L, r_L) is such that $\alpha \ln w_L + \beta \ln r_L \leqslant \alpha \ln w_k + \beta \ln r_k$ for all $k \in K$. We can also normalize this pair to $(w_L, r_L) = (1, 1)$, and, in this case, (w_k, r_k) show the ratios of wage and rent in location k to the lowest wage and rent combination.

threshold wage-rent combination, and any higher wage-rent combination that satisfies condition (9) and (10) can support the separation outcome. Nevertheless, as in traditional screening models, the wage-rent combination of (w_H, r_H) provides the only Pareto-undominated Nash equilibrium of this model.

Condition (9) provides several comparative statics results. An increase in the quality of H type increases the threshold wage and office rent for separation. An increase in the quality of contacts between L types reduces the separating threshold wage and office rent. In the same way, a reduction in the cost of maintaining a contact among the L type (i.e., decrease in c_L) reduces the separating threshold wage and office rent. A higher fixed cost of initiating a contact and learning about the quality of another agency leads to a larger separating threshold wage and office rent. Furthermore, a larger number of H type compared to L type agencies increases the threshold wage and office rent combination that is required for separation.

Pooling Equilibria—We defined and briefly discussed the pooling equilibrium in which all agencies choose to collocate at the lowest wage-rent location. We reject it on the grounds that an H type never wants to be close to L types as long as the number of low type agencies, N_L, is large and condition (7) is satisfied. Given (7), every contact with an L type decreases its overall productivity, and a large N_L is sufficient to deter H types from choosing the low wage-rent location.

Formally, a pooling outcome is not a Nash equilibrium if an H type agency prefers to deviate and stand alone. That condition simplifies to the following constraint

$$N_H \left((1-\eta) \left(\frac{\eta}{c_H} \right)^{\frac{\eta}{1-\eta}} q_H^2 - F \right) \leqslant N_L \left(F - (1-\eta) \left(\frac{\eta}{c_H} \right)^{\frac{\eta}{1-\eta}} q_H q_L \right). \tag{12}$$

This constraint implies that the benefits of networking with other H types cannot offset the adverse effect of being forced to participate in networking with L types.

To conclude, the most basic implication of the separation is that high quality agencies choose a location with high costs (i.e., wage-rent combination) that is not profitable for low quality agencies. In section 2, we provided evidence of segmentation and sorting in the location decisions of agencies that matches this prediction. In addition, the existence of a separating equilibrium requires that condition (8) be satisfied. We show in the next section that we can test whether this condition is or is not met.

4 Empirical Specifications and Estimation Methods

To derive our empirical specification, we start from (2) and transform it to log-linear form

$$\begin{aligned} \ln \Pi_{ik}^* &= a \ln A_0 + \ln N_k + \ln \left(T_0 c_i^{\frac{\eta}{1-\eta}} \right) + \ln \left(q_i \bar{q}_k + T_0 \frac{\rho}{N_k} c_i^{\frac{\eta}{1-\eta}} q_i - T_0 F c_i^{\frac{\eta}{1-\eta}} \right) \\ & \quad -\alpha \ln w_k - \beta \ln r_k + \varepsilon_{ik}, \end{aligned}$$

where $a = 1 - \alpha - \beta$, and $T_0 = \frac{1}{1-\eta}\eta^{-\frac{\eta}{1-\eta}}$.[32]

Using a Taylor expansion of $\ln\left(q_i \bar{q}_k + T_0 \frac{\rho}{N_k} c_i^{\frac{\eta}{1-\eta}} q_i - T_0 F c_i^{\frac{\eta}{1-\eta}}\right)$, one can show

$$\ln \Pi^*_{ik} = a \ln A_0 + \ln N_k + (\gamma_0 + \gamma_1 \ln q_i) \ln \bar{q}_k + \gamma_2 (\ln \bar{q}_k)^2 - \alpha \ln w_k - \beta \ln r_k + \mu_i + \varepsilon_{ik}, \tag{13}$$

where μ_i includes all the firm-specific terms in our Taylor expansion. We expect $\gamma_1 > 0$ and $\gamma_2 < 0$. In fact, we showed that $\gamma_1 < 0$ means that the necessary condition for separation is not satisfied and no separating equilibrium exists.[33] The sign of γ_0 is unclear and depends on the exact specification of the networking costs.

Thus, we have a discrete choice model with a variable coefficient on the average quality of agencies at each location. The variable coefficient is a function of agency i's quality. It increases as the quality of agency i increases. That is, the high quality agencies are more sensitive about their networks' quality than are the low quality agencies. This is exactly why the high quality agencies have an incentive to separate themselves from low quality agencies. In other words, we expect that the location choice of the high quality agencies makes it unprofitable for the low quality agencies to collocate with them. Put succinctly, $(\gamma_0 + \gamma_1 \ln q_i) \ln \bar{q}_k$ captures the quality matching by screening (separation) in our empirical model. Even though, based on the structural model, we strongly believe that is the only interactive term in our empirical model, we also examine a fully specified empirical model that includes all other interactive terms for the sake of completeness. Hereafter, we continue with a fully specified empirical model.[34] Furthermore, we break the $\ln\left(\bar{A}_{ik}\right)$ into firm and location specific effects. This provides us with the familiar error component model for the profits of agency i

$$\ln \Pi^*_{ik} = \sum_j \beta_j x_{kj} + \sum_{j,r} \gamma_{jr} z_{ir} x_{kj} + \mu_i + \delta_k + \varepsilon_{ik},$$

where μ_i, as before, collects all the additive firm-specific terms of the profit function and x_k is a vector of the location k's attributes, such as the logs of wage, rent, intra-industry scale or networking possibilities, number of broadcasting establishments, number of headquarters, and a measure of networks quality at the location (specifically, median employment of SU agencies). Moreover, z_i includes characteristics of agency i; namely, a measure of quality of agency i. Given equation (3), we use the median of employment at each location as the proxy for the overall measure of the quality of agencies at the location. The same equation

[32] Note that we multiplied all terms in a. Changes in the scale and intercept of the random profit do not affect the outcome of the discrete choice model (Train 2003).

[33] In Taylor series expansion, γ_0 is positive if and only if

$$\frac{\Delta \ln c}{\Delta \ln q} > \frac{1-\eta}{\eta} \frac{1}{\left(1 - \frac{\rho q_i}{N_k F}\right)}. \tag{14}$$

Assuming that $\frac{\rho q_i}{N_k F} < 1$, if the above condition holds, then the necessary condition for separation in (8) is satisfied. In this case, the separation is a likely outcome of the market and the main cause of segmentation in the advertising agency location decisions.

[34] Note that our favorite specification is nested in the fully specified model as all γs are zero, except γ_0 and γ_1 above.

also indicates that annual payroll of the movers at their origin can produce a good proxy for their quality (Jovanovic 1982). The fact that annual payroll of a mover is predetermined in its destination choice decisions makes payroll a reasonable exogenous measure of quality.

Following Berry, Levinsohn, and Pakes (2004) (henceforth, BLP04) and Bayer, McMillan, and Rueben (2004), we interpret the coefficients on terms $z_{ir}x_{kj}$ as how firms react based on their quality to the location attributes. We utilize the observed location choices by new firms (births) and the destinations of movers in the context of the above random profits model to estimate the parameters of the profit function, including the variable coefficients. We drop μ_i, because the profit maximizing location of agency i is unaffected by the invariant part of its profits over its choice set (Train 2003).

Thus, our empirical model is summarized in the following discrete choice model

$$\begin{aligned} y_{ik} &= \begin{cases} 1 & if \quad \pi_{ik} \geqslant \pi_{ij} \quad \forall j \in K \\ 0 & \qquad Otherwise \end{cases} \\ \pi_{ik} &= \ln \tilde{\Pi}^*_{ik} = \sum_j \beta_j x_{kj} + \sum_{j,r} \gamma_{jr} z_{ir} x_{kj} + \delta_k + \varepsilon_{ik} \end{aligned} \tag{15}$$

To entertain different assumptions about the parameters of the model and on the structure of the errors, we suggest four approaches to estimate this model. First, we ignore the heterogeneity among the agencies (i.e. $z_{ir} = \bar{z}_r$) and estimate the discrete choice model for births (new agencies) based on the location attributes alone. This results in a traditional Poisson regression model in subsection 4.1, where we assume there are no omitted attributes, and a GMM-IV estimation of an exponential model in subsection 4.2, that relaxes the assumption of the exogeneity of all covariates. Second, we investigate the possibility of heterogeneity/sorting by using firm characteristics in (15) and exploit a sample of movers to estimate the parameters with an emphasis on the coefficient of the interactive term, γ_1, that captures sorting and separation effects. This results in a conditional logit model in 4.3, if we assume that our covariates are exogenous to the location specific effects, δ_k, and a conditional logit model with location-specific fixed effects (BLP04) in 4.4, if we relax the exogeneity assumption.

4.1 Homogeneous Agencies and Exogenous Covariates

We start from the most restrictive set of assumptions, namely, we ignore the heterogeneity among the new agencies and we assume that there are no omitted location attributes that may be correlated with our covariates. Even though these may seem excessively strong assumptions, and we will relax them in the future, they are not without merit in a few respects. First, it is reasonable to think that a new agency is not fully aware of its potential and the period after the birth is the learning and self evaluation time (Jovanovic 1982; Duranton and Puga 2001). In this sense, the exact characteristics of a new agency are not fully known, even to its founders, prior to the birth. Second, these assumptions provides a specification and results that are comparable with estimates of wage and office rent elasticities, localization, urbanization, and local quality effects that are traditionally estimated using aggregate level variables.[35] Third, this provides the benchmark to compare with the results after controlling for the heterogeneity.

In this case, the model in (15) simplifies to

$$\pi_{ik} = \sum_j \tau_j x_{kj} + \delta + \varepsilon_{ik} \tag{16}$$

where $\tau_j = \beta_j + \sum_r \gamma_{rj} \bar{z}_r$ and $\bar{z}$ represents the common attributes shared by all new agencies. Given that ε_{ik} has an iid extreme value type-I distribution, it is easy to show that ML estimation of the parameters of this model is identical to ML estimation of the traditional Poisson regression model (Guimaraes, Figueirdo, and Woodward 2003; Chen and Kuo 2001; Cameron and Trivedi 1998) with

$$E\left[n_k \mid x_k\right] = \exp\left(\delta + \sum_j \tau_j x_{kj}\right) \tag{17}$$

where n_k is the number of births at location k.

This provides the benchmark model. We relax our assumptions in two directions: homogeneity of agencies and exogeneity of covariates in the following subsections.

4.2 Homogeneous Agencies and Endogenous Covariates

Here we take that the exponential specification in (17) is correct, but there are some unobserved location attributes, δ_k, which may be correlated with observed attributes. For example, recent revitalization projects in the downtown could increase rents for office spaces and also drawn agencies. Naturally, the omitted attributes are included in the residuals (error terms), thus the basic assumption that the covariates are uncorrelated with the error terms is violated.[36] Windmejier and Santos (1997) and Mullahy (1997) suggest a non-linear generalized IV estimation for the general exponential model based on the following set of moments

$$\sum_k \left(n_k - \exp\left(\sum_j \tau_j x_{kj}\right)\right) w_k = 0 \tag{18}$$

We discuss our list of instruments in the section 5.

[35] There is a caveat in interpreting the results. Although the statistical significance of the coefficients can be interpreted the same way as in linear regression models, the face value of the coefficients, like any other discrete choice model, are meaningless. This is because the coefficients can be arbitrarily scaled up or down and it would not affect the outcomes of the discrete choice model. The decision only depends on the relative values of the profit at different locations. In other words, the estimated profits are only an ordinal numbers. Nevertheless, the ratio of the localization coefficient to the wage coefficient, i.e. willingness to pay, is comparable with the same figure calculated from traditional models.

[36] Note that the ML first order conditions for Poisson regression are $\sum_k x_{kj} u_k = 0$ for all j, where $u_k \equiv n_k - \exp\left(\delta + \sum_j \tau_j x_{kj}\right)$. This is a sample analog of the moment conditions $E\left[u_k \mid x_k\right] = 0$. When there is no heterogeneity among the agencies, the above moment conditions are the same for all exponential models including the conditional logit.

4.3 Heterogeneous Agencies and Exogenous Covariates

At the beginning of this section, we show that agencies react differently to location attributes, specifically the quality of local networking opportunities, based on their own quality. We expect that the high quality agencies are more sensitive to the quality of local networks (that is $\gamma_1 > 0$). This is exactly why separation/sorting prevails.

In order to test this hypothesis, we need to estimate the parameters of (15) using the firm's attributes, specifically firm's quality. If our specification of profits captures all relevant location attributes (i.e. $\delta_k = \delta$), then we can easily estimate the parameters using the conditional logit model (McFadden 1974; Train 2003). Given (3), we would like to use the annual payroll as a proxy for agency quality. However, birth incidences do not provide the right kind of data to estimate this model. In fact, the choices of location and employment (size)[37] are likely to be simultaneous. Thus, we use the data on location decisions of movers instead of new agencies. The annual payroll of the mover in its previous location is predetermined to the relocation decisions.[38] That is, we estimate parameters of following profit function

$$\pi_{ik} = \sum_j \beta_j x_{kj} + \sum_{j,r} \gamma_{jr} z_{ir} x_{kj} + \varepsilon_{ik} \tag{19}$$

One attractive feature of this approach is that it is a direct extension of the Poisson regression in (4.1) for heterogeneous agencies.

4.4 Heterogeneous Agencies and Endogenous Covariates

Given that we use establishment level data to estimate (19), the covariates are not endogenous in the traditional sense, because the choice of a single establishment does not usually affect the attributes of a location (market level covariates). However, omitted location attributes can still result in correlation between observed attributes (e.g. rent, wage, or concentration) and error term. For example, wages may be higher for the location with desirable unobserved attributes for businesses, such as loose environmental regulations. Following Berry, Levinsohn, and Pakes (2004),[39] we employ a set of location-specific dummies

[37] As a matter of fact, almost all characteristics of a new agency are decided simultaneously. We have thought of some predetermined measures of quality for new agencies. The most interesting possibility is to use the previous experience of the founders of a new agency. We are still working on this.

[38] In fact, this follows the logic in the subsection (4.1). As the agencies learn about their quality, they may find the current location a poor match for their quality and decide to move to a different location based on their new understanding. Nevertheless, their annual payroll (size) at the current location partially reflects their quality.

[39] Note that their estimation procedure is more complicated than the one we use here. It is because car prices are endogenous to their identification of demand for cars. Thus, they need an addition step to calculate a set of equilibrium prices.

(fixed effects) to absorb all the unobserved location attributes. That is,

$$\pi_{ik} = \lambda_k + \sum_{j,r} \gamma_{jr} z_{ir} x_{kj} + \varepsilon_{ik} \tag{20}$$

$$\lambda_k = \sum_{j} \beta_j x_{kj} + \delta_k \tag{21}$$

We estimate (20) using the ML method for the condition logit model. This produces an unbiased estimate of γ_1, the coefficient for the quality interactive term, in (13) (and other γs in the fully specified model). To estimate the overall elasticities, we use $\hat{\lambda}_k$ and estimate (21) by 2SLS.

5 Data

We use the 1997, 1992, and 1977 Standard Statistical Establishments List (SSEL). These contain the universe of all establishments in the US in those years. Those include the social security (tax records) records of all establishments. They are augmented by the Census Bureau to cover all establishments in the US (using other sources, including previous Economic Census information, organization surveys, and data from other federal agencies) as the base mailing lists used by the Census Bureau to conduct the Economic Census. They are then corrected and populated using the responses.[40] The first benefit of using the SSEL is that it includes all the establishments. Thus, unlike all other publicly available data, it is not prone to selection biases and sampling shortcomings. Second, the large number of observations allows us to use finer geographical units than previously used (detailed county level data rather than state level data used in King et al. 2003). Third, the reported values are more reliable than other data because they are from tax records and are corrected and augmented (for example to include all establishments of multi-unit firms) during the Economic Census. For this very reason, we also limit ourselves to the Census years.

Using the SSEL for 1997 and 1992, we construct the list of all new SU and MU advertising agencies (births) in the continental US (excluding Alaska, Hawaii, and Puerto Rico). A birth is a Census File Number (CFN) in 1997 that did not exist in 1992. The stocks of SU and MU agencies are also constructed in 1992 and 1977. We use a consistent set of 3072 US counties over time to generate aggregate data at the county level. We focus on 812 urban counties (1994 definitions) because more than 95 percent of all agencies and 99 percent of all employment in advertising agencies were located within those urban counties in 1997. In other words, virtually the universe of advertising agencies is the set of urban counties. There were more than 11,000 SU births and more than 600 MU births in 1992-97 (Table 4). In the same period, 403 single unit agencies moved into these 812 urban counties from other locations in the continental US (only 188 urban

[40]The response to the Economic Census is required by law (Title 13, United States Code). The same law ensures that the Census report is confidential.

counties received some movers).[41] However all the covariates, including their own (non-zero) employment, are recorded for 340 of the movers. Among these movers, 166 moved into 76 CBDs and the rest moved into 95 suburban counties. The importance of the sample of movers is that we can treat their attributes in their previous location as predetermined to the relocation decisions.

The single largest expenditure in the advertising agency business is payroll. About 72 percent of the operational costs of agencies are salaries and other fringe benefits paid to employees. This emphasizes the importance of wages in the profit function of advertising agencies. We construct the wages in the industry in 1992 using the median salaries of establishments for each PMSA from the Census Bureau's newly released Longitudinal Business Database (LBD).[42] We think that PMSAs provide reasonable boundaries for local labor markets rather than counties or Census regions.

The office rent data for 1992 is from *Comparative Statistics of Industrial and Office Real Estate Market (1993)*, the annual publication of Society of Industrial and Office Realtors. This publication reports the rent for Class "A" office space in 111 US markets in 1992 (131 in 2003). The rents are reported for CBDs and suburbs. The office space rents for 106 CBDs are constructed by matching this data to the set of CBDs from the 1992 SSEL, and another 329 counties are assigned suburban office rents by matching the corresponding PMSAs. Thus, the office rents are more reliable and have more variation for CBDs than for all urban counties. Therefore, we estimate the parameter of the profit function using the birth incidences in 106 CBDs with available office rents.[43] In addition, we estimate our model using the data for 211 CBDs that have all the variables (including the wages and all the instruments), using total employment and a multi-county PMSA dummy to proxy the effect of office rent (based on the traditional rent gradient story (see Henderson 1988)); and also for 659 urban counties, using CBD/suburb dummies in addition to the above variables. For the movers, we estimate the parameter using all urban counties with a non-zero number of moves. This provides us with 161 counties (10 counties out of 171 urban counties with some moves miss one or more of the other covariates and/or instruments).

We include covariates to capture localization (within industry networking) effects (Henderson 2003), backward and forward linkages and local demand (Porter 1990), and urbanization externalities (Dixit and Stiglitz 1977; Either 1981). In so doing, we construct the establishment stocks of SU agencies for each of 812 urban counties using 1992 SSEL. We also calculated the stocks of establishments and employments in radio and television broadcasting (SIC-483), newspapers and periodicals (SIC-271 and 272), and graphic design

[41] A move is a change in the county FIPS codes in the two consecutive Economic Census (here, 1992 and 1997) for an existing CFN. This may underestimate the real number of moves, since some interstate moves may coincide with changes in CFN numbers due to changes in tax division and therefore changes in Employment Identification Numbers (EIN).

[42] The LBD is put together by the Census Bureau-CES using the SSEL for all businesses (Jarmin and Miranda 2002), to mirror the well-known LRD.

[43] The results for 435 urban counties with reported office rents, including the sample of 106 CBDs, are also available. They are very similar to the CBDs' results.

(SIC-7336) using the LBD; and in headquarters using the Central Administration Offices (CAO) database.

In terms of instruments, we use a variety of topographical, historical, and socioeconomic variables. Our goal is to disentangle the effects of our main variables—namely, rents, wages, stocks of agencies (networking and localization), network qualities, and linkages to local demanders and suppliers—from other local unobservables (including local shocks). In general, we worry that unobservables may affect some of the covariates; and we are, specifically, worried about the endogeneity of stock variables. As we discussed in the last section, we use the instruments in two specifications to correct the possible biases introduced by the above difficulties. First, we use the instrumental variables in the generalized method of moments estimate for the Poisson (more accurately, generalized exponential model) specification of profits function using births at county level (Windmeijer and Santos 2002). Second, we use our instruments in instrumental variables estimate for county-specific fixed effects on county level covariates (BLP04 method).

Our socioeconomic instruments are mostly drawn from the 1970 Census of Population and Housing. They include several measures of housing availability and quality at the county level, such as the existence of large housing complexes (share of housing units consisting of 10 or more units), the ratio of housing units to business establishments, and the median residential rent for one bedroom apartments in 1990. We expect that these variables are affected by various local and regional laws and regulations that do not directly affect the advertising agency business and also are most probably uncorrelated with local shocks in the 90s. However, these variables can be seen as amenities (or disamenities) to the local labor force and therefore affect wages (Roback 1982).[44] They also may present certain shortages or abundances in the local office space market and affect office rents.

In addition, we use local government expenditures in 1987 at the county level as a measure of overall public services that are provided. We expect that it has positive effects on rents, both because it requires higher property taxes (Henderson 1985; Oates 1969; Tiebout 1956) and because government expenditures could have a crowd out effect on the office real estate market. The effect on wages is uncertain, since on the one hand we expect that it will have negative effects on wages as long as local public services are considered positive amenities (Roback 1982), but on the other hand the government expenditures drive wages higher due to increase in the local labor demand.[45] Furthermore, local government expenditures my increase the demand for advertising as some local government use media for public awareness programs (for example anti-smoking campaign in Massachusetts). In this sense, it provides an instrument for the stocks of agencies in 1992.

Our historical variables are from the U.S. Dept. of Commerce, Bureau of the Census (2001) (corrected

[44] Though we expect that our historical wage variables will provide a stronger instrument for wages in 1992.

[45] In basic urban models (Henderson 1988), higher labor demand is relieved by in-migration to the city. Wages rise as the city size increases.

by Haines). We use the market potential and retail wages in 1930. We hope that these pick up the original location of the advertising industry at the beginning of the 20th century. The changes in advertising since that time have been so great that it is impossible to think of any correlation of these measures with the current shocks (or even shocks in the last 30 years) to the industry. The industry today not only uses different kinds of media, but also it performs different activities (see section 2). We also use a single geographical dummy that indicates if the county is 80 miles from an ocean coast line (Rappaport and Sachs 2003).

In addition, we use a few historical industry-specific variables to instrument for wages, local scale and networking, and the quality of local networks. We think of median wages in 1977 as a reasonable instrument for wages in 1992 and the median size of agencies at the county level in 1977 as an instrument for local network quality in 1992. The main justification for using this variable is that we know that the geographical patterns of the stocks and births of single unit agencies have changed over 1977-1992 (see section 2). This rules out the possibility that both our 1977 (instruments) and 1992 variables (covariates) are mainly driven by the same set of long-lasting local unobservables. That would have made both set of variables correlated with long lasting location unobservable attributes (buried in residuals). Rather, the only worry is possibility of correlation among the shocks in 1977 and 1992. However, we believe that the unobserved shocks to the locations (CBDs) attributes (amenities and/or disamenities) were different in 1977 and 1992, because the business environment have changed since 1970s. Specifically, in advertising, the demise of the old commission system in the late 1980s caused shifts in the location of agencies since 1977 (usually through new births). Therefore, we run simple experiments to test for this possibility. We include the instruments in the ordinary Poisson specification, and we expect the joint effect of these variables to be insignificant and the coefficients on our base variables to be unaffected. This would indicate that our instruments are unlikely to be correlated with the current local shocks and possible omitted local attributes. Finally, we use the Sargan test for overidentifying restrictions to test if our instruments are valid, assuming that the model is well specified.

Even though we argue that both 1992 and 1977 variables are not driven by the same set of long-lasting local attributes (observable ro not) and that the local shocks in 1992 are different from 1977, the characteristics of agencies in 1992 still partially reflect that of 1977 because of the coordination failure among the network members to relocate together, the persistence of agglomeration effect, and the deterrence of general moving costs.[46] Hence, we feel that historical attributes of SU advertising agencies could provide good instruments for their current counterparts.

[46]Remnant agencies provide enough correlation between historical and current attributes of local SU agencies.

6 Results

6.1 Homogeneous Agencies

Following our discussion in section 4, Table 3 reports the estimates of the models in subsection 4.1 for different sets of covariates. Our dependent variable is the birth counts of single unit agencies during 1992-97 at the county level. We match the birth counts to the county attributes for 1992. As a result, we construct two samples. The first sample contains 106 CBDs for which we were able to find corresponding rents for class "A" office spaces. All our covariates and instruments are available for this sample. The second sample includes 211 CBDs with no missing values for the covariates except office rents.[47] For the second sample we proxy the effect of rents with two variables, a dummy for multi-county PMSAs and the total employment of the county in 1992. In this respect, we follow standard urban models, which state that the rent gradient, and therefore rent at the CBD, is a rising function of city size (Henderson 1988). The above two variables are intended to capture the city size. This doubles the sample size from 106 CBDs to 211 CBDs and expands our sample to include most of the medium sized cities. The increase in the sample size reinforces our confidence in our results in three ways. First, it confirms that our results are not driven by our specific sample of 106 CBDs (mostly the CBDs of larger cities). Second, it allows us to go beyond estimating the 5 or 6 parameters of the base model and have meaningful estimates of coefficients for broadcasters and headquarters. Of course, we could use all the urban counties to further expand our sample; however, we do not feel it is appropriate to compare CBD with non-CBD counties in the absence of county-specific fixed effects. Third, this specification (without rent) provides comparable results to our heterogeneous agency results in the next subsection where we employ the sample of movers.

In addition to the log of rent for class "A" office space (in dollars per square foot), other main covariates are the log of wage in the advertising industry (median wage from LBD), the local scale of advertising agencies (captured by the log of SU counts), the log of median SU size in 1992, and its square.

[insert Table 3]

The results in columns I-III show large own scale effects. The coefficients on the log of SU counts (own scale elasticity) are about 1 and highly significant in all the specifications. The rent and wage elasticities have the right signs and their relative magnitudes correspond to our expectations, given that the worker compensation is the largest item in advertising agency expenditures.[48] Interestingly, the coefficient on agency size (the proxy for quality) is negative and significant in columns I, III, IV, and VI, when we do not include

[47] We experimented with a third sample of all urban counties (CBD and non-CBD). All covariates (except office rent) exist for 659 out of 812 urban counties. We add a dummy for CBD counties in order to control for fixed differences between CBD and non-CBD counties. The results are not significantly different from those of our 211 CBD sample.

[48] As we point out in section 2, worker compensation accounts for 72 percent of agency business expenses, and the share of rent and lease is less than 4 percent of expenses.

its square. We think of four reasons to obtain a negative sign for coefficient of average quality. First, our model only predicts a negative sign for the squared term and is silent about the sign of the linear term. As we include the square term, the coefficient of the linear term becomes insignificant in column II, and even for the larger sample, it becomes positive, but still insignificant. This suggests that there may not be enough variation in the SU size among CBDs to help us separate the effects of size (quality) from its square. Second, in the absence of the interactive term, which captures the sorting effects, the coefficient on size and its square are most certainly biased. Third, in general, all the estimated coefficients are likely to be biased due to omitted county attributes and possible endogeneity. Fourth, the fact that the higher quality agencies are fiercer competition—something that we have not modelled—may cause an originally negative effect from higher quality neighbors. However, we expect that correcting for the matching on quality will dampen this effect and show positive overall effects on neighbors' quality for higher quality agencies. This is precisely what we will see in the next subsection.

We tackle the problem of omitted county attributes (and possible endogeneity) using instrumental variable estimation for the general exponential model, explained in 4.2.[49] Table 4 shows the results. In general, the magnitude of the wage elasticity increases and the effect of rent disappears. The own elasticity scale (localization effect) also rises.[50] The coefficient on the quality of agencies in the location (networking quality) is insignificant. In some cases, especially for the smaller sample, the estimation procedure does not converge when we add the square of our local quality measure. However, whenever it converges, the coefficient on the squared quality (squared log of median size) is insignificant too. As we discussed above, we expect controlling for heterogeneity and sorting among agencies will sharpen the results.

Table 4 shows that our specification and instruments pass the Sargan overidentifying restrictions test with a wide margin, lending confidence to our choice of instruments and the instrumental variable estimation results. Nevertheless, we use additional experiments to test our instruments.[51]

[insert Table 4]

6.2 Heterogeneous Agencies

Now we move to the main set of estimation results where we control for heterogeneity among agencies. We use the individual observations for the SU agencies that relocated during the 1992-97 period. We use their

[49] We use the Gauss program EXPEND (Windmeijer 2002) to preform the GMM-IV estimation of the general exponential model. This program provides heteroscedasticity corrected standard deviations using the White sandwich formula for the non-linear IV estimator.

[50] The estimates of the own scale elasticities (localization effects) are extremely large in all of our estimates. Increase in this elasticity is not what we expected.

[51] In an informal exercise, we include our instruments in our base Poisson model for births. The idea is that if the instruments are correlated with unobservables, then their coefficients should be significant in the Poisson regressions. Rejecting the significance increases our confidence in the instrument. We do not find evidence to contrary for our controversial instruments, such as median agency size (proxy for quality) in 1977, wage in advertising in 1977, and local government expenditures in 1987. However, we obtain effects from some of the less controversial ones (in our opinion), such as market potential in 1930.

size while in operation at their original location (annual payroll in 1992) as a proxy for their quality. We can now estimate γ_1 in equation (13) and test the separation hypothesis. Since we have a limited number of movers, we are forced to use all 340 moves including the moves into non-CBD urban counties. The sample is evenly divided, with 76 CBD counties and 95 non-CBD urban counties (see Table 2 and section 5 for information on the sample of movers). We include an additional dummy to capture the differences between the CBD and non-CBD (suburban) counties; however, we are not worried about this because we are able to use county-specific fixed effects to get an unbiased estimate of γ_1, the coefficient of main interest in our model.

Table 5 shows that γ_1 is unequivocally positive and highly significant. Even its magnitude does not change with the introduction of other interactive terms. The main set of results are in columns IV-VI. In these estimates, in the first stage we include a set of county-specific fixed effects in addition to interactive term(s). This specification provides unbiased estimates of the coefficient of the interactive terms, because all the unobserved county-specific attributes are absorbed by the fixed effects. These results are reported in the upper right section of the table. The rest of the coefficients in the lower right corner of the table are estimated with 2SLS using the county-specific fixed effects as the dependent variable and county attributes as the covariates.

[insert Table 5]

Columns I-III show the results for the ordinary conditional logit including the variable coefficient on networking quality, which varies with agency quality. Column II provides an estimate of our exact structural model. The own scale elasticity is large and positive. The coefficient on the interaction between agency quality and local network quality is positive and highly significant. In addition, the coefficient on the square of network quality is negative and significant at a 10 percent confidence. The only issue is the non-negative but insignificant wage elasticity. In column III, we estimate the fully specified model that includes all the interactions of agency quality with all the location attributes. In other words, we use variable coefficients on all the county attributes. The wage elasticity is negative and significant. The benefits of network quantity (SU median size) steeply rise with agency quality.[52] The overall effect of network quality becomes positive as our proxy for agency quality (log of annual payroll) is larger than 3.7. Considering that the average quality (log of annual payroll) of movers is 4.14, the overall effect of network quality is positive for an average mover.

As we discuss in subsection 4.3, the estimated parameters of the ordinary conditional logit specification may, nevertheless, be biased if the unobserved location attributes are correlated with covariates. Therefore we employ a set of county-specific dummies to absorb the contributions of all the county attributes (observed and unobserved) and follow the procedure in subsection 4.4. The results are reported in columns IV-VI.

[52] The coefficient reported in the first line of Column III shows the slope of the variable coefficient on local network quality. It is 0.150 and significant.

We are able to identify 166 county-specific fixed effects. In column V, we estimate the exact specification in equation (13) (our structural model). The results indicate expected signs for all coefficients, except for wage coefficient that is insignificant; however, the coefficient on both local network quality and its square are insignificant.[53] In column IV, we only include the linear effect of local network quality, and the coefficient is negative and significant. Therefore, the effect of local network quality is negative for low quality agencies (competition), but it turns positive and large for high quality agencies.

Finally, in column VI, we estimate the fully specified model including all interactive terms.[54] The expected sign appears for the wage elasticity. The quality matching coefficient, γ_1, increases to 0.201. The overall positive effect of local network quality appears for agencies of size larger than 4.5. This means that the elasticity (value) of network quality for an average agency in New York is about 1.8.[55] In other words, the profit of an average agency in New York (the location with the highest quality agencies) increases by 17 percent if the average local network quality, measured by average employment, increases by 1 employee.

7 Conclusions

We show that segmentation and sorting are dominant features in the location decisions of advertising agencies. Even though these estimates are limited to advertising agencies, we think that they are representative of similar effects for many business services.

There are several directions that we can take in future research to improve upon the structural model and empirical approach. Introducing a continuum of quality among agencies would result in a richer structural model that not only allows further understanding of market segmentation, sorting, and geographical concentration within service industries, but also provides a new perspective on the effects of the business environment (such as corporate tax and labor regulations) on establishments' productivity and location decisions. Regarding our empirical specification, a need exists for stronger instruments, specifically for the agency quality. In addition, we think that for movers, we can do a better job of capturing their (predetermined) quality by using an iterative method that uses the estimated parameters of a profit function to re-estimate the quality of a mover in its previous location.

[53] We think that it is because we are on the highly linear declining portion of the quadratic function. Thus, it is difficult to estimate the linear and quadratic effects together.

[54] We report a Hansen-Sargan J-test for overidentifying restrictions at the bottom of columns IV-VI. The test cannot reject the validity of our instruments even at 10 percent confidence level. The margin is much higher for columns IV and VI.
In addition, we experiment with the residuals of 2SLS estimation and interact them with agency quality and include this interactive term in our original conditional logit step. We are interested to see if the coefficient of this interactive term is significant or not. A significant coefficient would raise a red flag in the sense that some of our unobserved county attributes may effect the location decisions of movers beyond the location fixed effects. In other words, different quality agencies may react differently to those unobservable attributes. We obtain an insignificant coefficient for the column VI specification. We obtain mixed results for column IV based on the set of instruments.

[55] Note that the average annual payroll of SU agency in 1997, from Table 1, is $\ln(10.7 \times 61200) = 13.4$.

8 References

Arzaghi, Mohammad (2005), "A Model of Knowledge Sharing and Network Formation among Advertising Agencies," Working Paper, Brown University.

Arzaghi, Mohammad; Henderson, J. Vernon (2005), "Networking Off Madison Avenue," Working Paper, Brown University.

Bagwell, Kyle (2001), *The Economics of Advertising*, Edward Elgar Publishing, Inc.

Bagwell, Kyle (2003), The Economic Analysis of Advertising, Work in Progress, March 2003.

Bala, Venkatesh; Goyal, Sanjeev (1998), "Learning from Neighbours," Review of Economic Studies, 65, 595-621.

Bayer, Patrick; McMillan, Robert; Rueben, Kim (2004), "Residential Segregation in General Equilibrium," Working Paper.

Berry, Steven; Levinsohn, James; Pakes, Ariel (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," JPE, 112(1), 68-105.

Cameron, Colin; Trivedi, Pravin K. (1998), *Regression Analysis of Count Data,* Cambridge: Cambridge University Press.

Chen, Zhen; Kuo, Lynn (2001), "A Note on the Estimation of the Multinomial Model With Random Effects," American Statistician, 55(2), 89-05.

Clark, Erick (1990), *The Want Makers: The World of Advertising,* New York: Viking.

Coase, Ronald (1960), "The Problem of Social Cost," Journal of Law and Economics, 3(1), 1-44.

Davis, James C.; Henderson, J. Vernon (2004), "The Agglomeration of Headquarters," Woring Paper, Brown University.

Dixit, Avinash; Stiglitz, Joseph (1977), "Monopolistic Competition and Optimum Product Diversity," AER, 67(3), 297-308

Duranton, Gilles; Puga, Diego (2001), "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products," AER, 91(5), 1454-1477.

Ellison, Glenn (1993), "Learning, Local Interaction and Coordination," Econometrica, 60(5), 1047-1071.

Ellison, Glenn; Fudenberg, Drew (1993), "Rules of Thumb for Social Learning," JPE, 101(4), 612-643.

Epple, Dennis; Sieg, Holger (1999), "Estimating Equilibrium Models of Local Jurisdictions," JPE, 107(4), 645-681.

Fujita, Masahisa; Ogawa, Hideaki (1982), "Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations," RSUE, 12(2), 161-196.

Gladwell, Malcolm (2002), *The Tipping Point: How Little Things Can Make a Big Difference,* Boston: Back Bay Books.

Glaeser, Edward; Gyourko, Joseph (2002), "The Impact of Zoning on Housing Affordability," NBER Working Paper 8835.

Guimarães, Paulo; Figueirdo, Octávio; Woodward, Douglas (2003), "A Tractable Approach to the Firm

Location Decision Problem," Review of Economics & Statistics, 85(1), 201-204.

Hameroff, Eugene J. (1998), *The Advertising Agency Business*, New York: McGraw-Hill.

Henderson, J. Vernon (2004), "What Makes Big Cities Tick? A Look at New York," Working Paper.

Henderson, J. Vernon (2003), "Marshall's Scale Economies," Journal of Urban Economics, 53, 1-28.

Henderson, J. Vernon (1997), "Medium Size Cities," Regional Science and Urban Economics, 27, 583-612.

Henderson, J. Vernon (1985), "Property Tax Incidence with a Public Sector," JPE, 93(4), 648-665.

Jackson, Matthew; Wolinsky, Asher (1996), "A Strategic Model of Social and Economic Networks," Journal of Economic Theory, 71, 44-74.

Jaffe, Adam B.; Trajtenberg, Manuel; Henderson, Rebecca (1993) "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," QJE, 108, 557-598.

Jarmin, Ron S.; Miranda, Javier (2002), "The Longitudinal Business Database," Washington DC: Center for Economic Studies, U.S. Bureau of the Census, Working Paper, No. CES-WP-02-17.

Jovanovic, Boyan (1982), "Selection and the Evolution of Industry," Econometrica, 50(3) 649-670.

King III, Charles; Silk, Alwin J.; and Ketelhohn, Niels (2003), "Knowledge Spillovers and Growth in the Disagglomeration of the US Advertising-Agency Industry," Journal of Economics and Management Strategy, 12(3), 327-362.

Maskin, Eric S. (1994), "The Invisible Hand and Externalities," AER, 84(2), 333-337.

McDonough, John et al. (2003), *The Advertising Age Encyclopedia of Advertising,* New York: Fitzroy Dearborn.

Mullahy, John (1997), "Instrumental-Variable Estimation of Count Data Model: Applications to Model of Cigarette Smoking Behavior," The Review of Economics and Statistics, 79(4), 586-593.

Nicosia, Franesco M. (1971), *Advertising, Management, and Society,* New York: McGraw-Hill.

Petrin, Amil; Train, Kenneth (2000), "Omitted Product Attributes in Discrete Choice Models National Bureau of Economic Research, Inc, NBER Working Papers: 9452.

Rappaport, Jordan; Sachs, Jeffery (2003), "The United States as a Coastal Nation," Journal of Economic Growth, 8(1), 5-46.

Ray, Debraj; Vohra, Rajiv (2001), "Coalitional Power and Public Goods," JPE, 109(6), 1355-1384.

Riley, John G. (1979a), "Information Equilibrium," Econometrica, 47, 331-360.

Riley, John G. (1979b), "Testing the Educational Screening Hypothesis," JPE, 87(5), 227-252.

Riley, John G. (1976), "Information, Screening and Human Capital," AER, 66(2), 254-260.

Roback, Jennifer (1982), "Wages, Rents, and the Quality of Life," JPE, 90, 1257-78.

Rosenthal, Stuart; Strange, William (2003), "Geography, Industrial Organization and Agglomeration," Review of Economics and Statistics, 85, 178-188.

Rothschild, Michael; Stiglitz, Joseph (1976), "Equilibrium in Competitive Insurance Market: An Essay on the Economics of Imperfect Information," QJE, 90(4), 629-650.

Schmalensee, Richard (1972), *The Economics of Advertising*, Amsterdam: North-Holland Publishing Company.

Schmalensee, Richard; Silk, Alvin J.; Bojanek, Robert (1983), "The Impact of Scale and Media Mix on Advertising Agency Costs," Journal of Business, 56(4), 453-475.

Shy, Oz (1995), *Industrial Organization: Theory and Applications*, Cambridge, Massachusetts: MIT Press.

Silk, Alvin J.; Berndt, Ernst R. (1994), "Costs, Institutional Mobility Barriers, and Market Structure: Advertising Agencies as Multiproduct Firms," Journal of Economics and Management Strategy, 3(3), 437-450.

Stiglitz, Joseph; Weiss, Andrew (1983), "Screening Alternative Approaches to Analyzing Markets with Asymmetric Information: Reply," AER, 73(1), 246-249.

Society of Industrial and Office REALTORS; Landauer Associates, Inc. (1993), *Comparative Statistics of Industrial and Office Real Estate Market,* Washington, DC: Society of Industrial and Office REALTORS.

Tirole, Jean (1988), *The Theory of Industrial Organization*, Cambridge, Massachusetts: MIT Press.

Train, Kenneth (2003), *Discrete Choice Methods with Simulation,* Cambridge, MA: Cambridge University Press.

U.S. Dept. of Commerce, Bureau of the Census (1992), Decennial Census reports, 1890-1960 [investigator]. Historical, Demographic, Economic, and Social Data: The United States, 1790-1970, ICPSR Study No. 3 [computer file]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor]. Corrected and amended by Michael R. Haines, Department of Economics, Colgate University, 2001.

Wells, William D.; Burnett, John; Moriarty, Sandra (2003), *Advertising: Principles and Practice (6th Edition),* New Jersey: Prentice Hall.

Windmeijer, Frank (2002), "EXPEND, A Gauss Program for Non-Linear GMM Estimation for Exponential Models with Endogenous Regressors for Cross Section and Panel (Dynamic) Count Data Models," http://cemmap.ifs.org.uk/docs/cwp1402.pdf.

Windmeijer, Frank; Santos Silva, Joao M.C. (1997), "Endogeneity in Count Data Model: An Application to Demand for Health Care," Journal of Applied Econometrics, 12(3), 281-294.

Wilson, Charles A. (1977), "Education and Screening," AER, 67(5), 949-958.

Yabushita, Shiro (1983), "Theory of Screening and the Behavior of the Firm: Comment," AER, 73(1), 242-245.

Yi, Sang-Seung (1997), "Stable Coalition Structures with Externalities," Games and Economic Behavior, 20, 201-237.

Figure 1: Location quotients of advertising establishments against the log of total employment for all CBDs in the US.

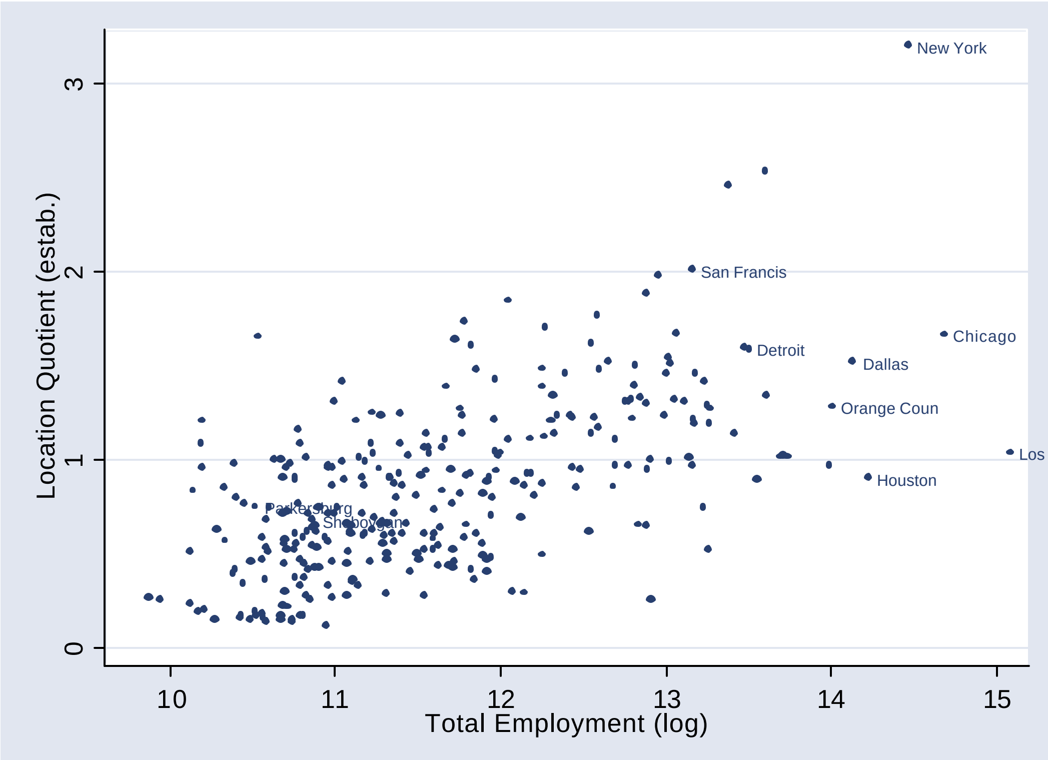

Figure 2: Agency sizes and agglomeration of advertising, 1997.

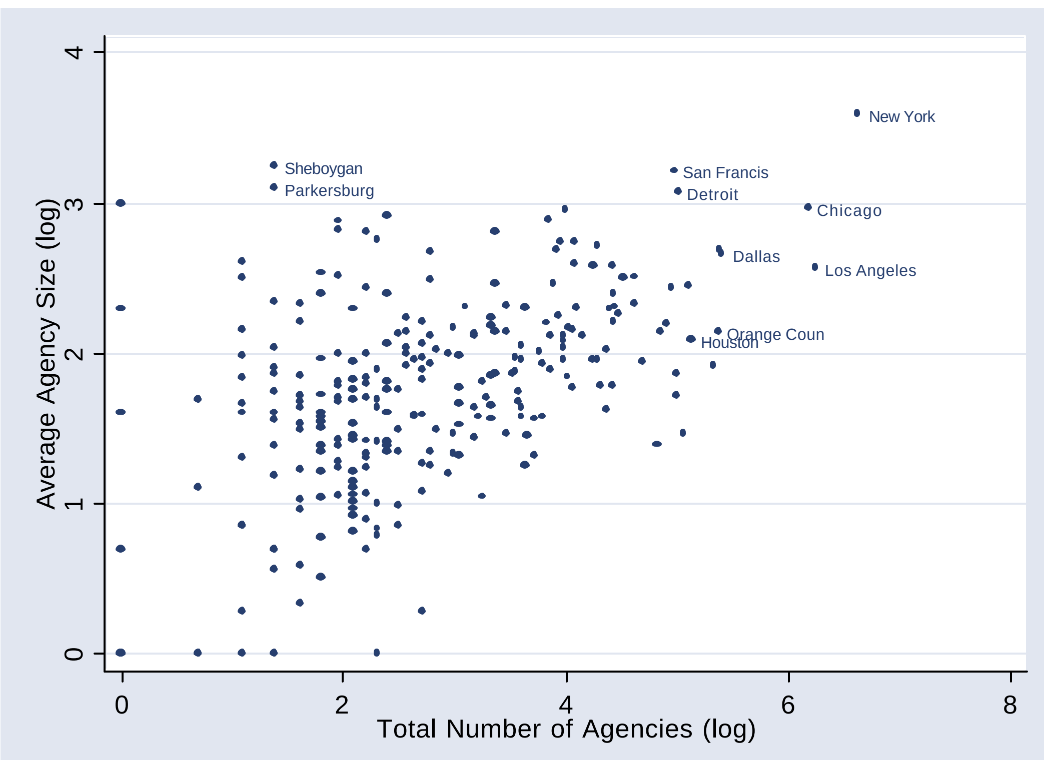

Figure 3: Agency sizes and cost of office space (dollars per square foot), 1997.

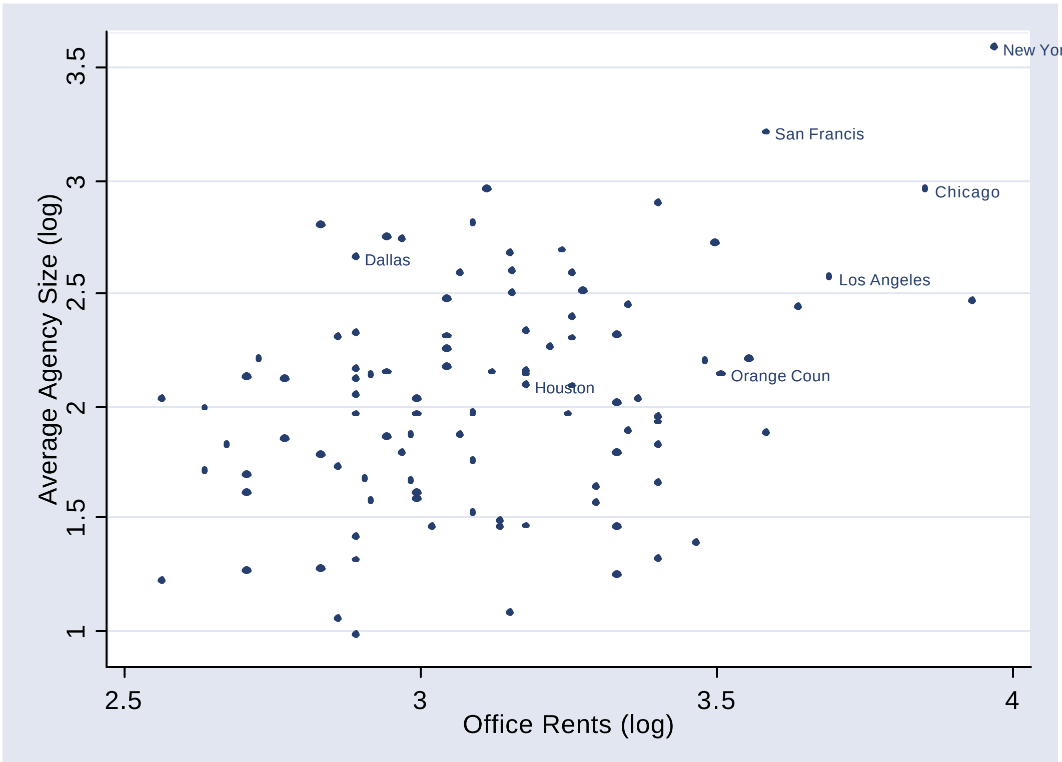

Table 1: Characteristics of single and multi-unit agencies in the United States and the top three advertising centers (CBDs), 1997.

	United States[1]		New York	Chicago	Los Angles	Others
	MU[2]	SU	SU	SU	SU	SU
Establishments	1,166	12,137	626	417	457	10,637
Average sales/receipts ($000)	7,280	685	1,420	810	1,010	623
Average employment	54.4	6.2	10.7	6.6	7.0	5.9
Average salaries ($)	53,300	38,900	61,200	48,900	47,600	36,900
Rental rate, class "A" office space ($/sq. ft.)[3]	23.6		52.8	47.0	40.0	23.0
Share of incomes from media commissions	47.8	27.6	39.6	28.2	38.2	24.4
Share of incomes from commissions on materials and production services	8.7	18.6	20.6	24.9	9.5	18.6

Sources: Census Bureau, 1997 Economics Census, Census of Services; SIOR, Comparative Statistics of Industrial and Office Real Estate Markets, 1992.
The averages are based on the responses to the 1997 Census of Services and include a subset of all establishments.
[1] The figures for the U.S. exclude Alaska, Hawaii, and Puerto Rico.
[2] These establishments are organized into 387 firms.
[3] Average quoted gross rental rate for class "A" competitive office space in U.S. dollars per square foot.

Table 1': Aggregate characteristics of advertising agencies in the United States and the top three PMSAs, 1997.

NAICS-54181	United States	New York	Chicago	Los Angeles
Establishment	13,390	914	480	516
Receipts[1]	16,872	4,085	1,370	943
Media Billings[1]	70,620	20,285	6,344	4,799
Billings on Materials and Services[1]	18,441	2,576	1,205	513
Annual Payroll[1]	7,557	2,055	616	417
Employment	139,486	27,646	9,137	6,751
Emp. per Establishment	10.4	30.2	19.0	13.1
Receipt per Establishment	1,260,045	4,469,365	2,854,167	1,827,519
Receipt per Employee	120,958	147,761	149,940	139,683
Average Salary	54,177	74,333	67,418	61,769
Receipts in Professional and Scientific Services[1]	579,542	43,494	30,849	31,679
Advertising Share	2.9	9.4	4.4	3.0

Source: All figures are from Census of Service Industries, Geographical Area Series, if it is not specified otherwise.
1. The figures are in million of dollars.
2. The billings figures are adjusted from the State figures to PMSA level based on the total receipts ratios.

Table 2: Description of the single unit (SU) births and movers samples, 1992-97.

SU agencies	CBDs	Suburbs	Total
Briths 1992-97	8,618	3,151	11,769
Counties with at least one birth	305	333	638
Movers 1992-97	166	174	340
Counties with at least one move	76	95	171

Sources: Census Bureau, SSEL 1992 and 1997.

Table 3: Estimating the parameters of pr ofit function using the location decisions of new single unit (SU) agencies and Poisson regression model, 1992-97.

Births (Poisson Regressions)	**Central Business Districts with Rent**			**All Central Business Districts**		
	(I)	**(II)**	**(III)**	**(IV)**	**(V)**	**(VI)**
Office Rent	**-0.140**	**-0.140**	**-0.167**	..	..	..
	(0.054)	(0.054)	(0.058)			
Multi-county PMSA	..	..	..	**-0.118**	**-0.122**	**-0.118**
				(0.025)	(0.025)	(0.026)
Employment 1992	..	..	..	0.030	0.025	-0.025
				(0.039)	(0.039)	(0.057)
Wage	**-0.456**	**-0.452**	**-0.343**	**-0.318**	**-0.316**	**-0.266**
	(0.075)	(0.076)	(0.079)	(0.064)	(0.064)	(0.066)
SU scale	**1.090**	**1.090**	**1.090**	**1.026**	**1.027**	**1.003**
	(0.020)	(0.020)	(0.042)	(0.035)	(0.035)	(0.036)
SU size (median)	**-0.234**	-0.493	**-0.247**	**-0.221**	0.065	**-0.255**
	(0.055)	(0.392)	(0.062)	(0.047)	(0.268)	(0.049)
SU size squared	..	0.106	..	..	-0.110	..
		(0.158)			(0.102)	
Broadcasting	..	..	**0.118**	..	..	**0.089**
			(0.030)			(0.029)
Headquarters	..	..	**-0.101**	..	..	-0.010
			(0.033)			(0.037)
Total establishments	..	..	..	..	..	..
Obs.	106	106	106	211	211	211
Pseudo R2	0.878	0.879	0.881	0.884	0.884	0.885

Coefficients in bold face are significant at 5 percent and in bold-italic are significant at 10 percent.

Table 4: Estimating the parameters of profit function using the location decisions of new single unit (SU) agencies and GMM-IV method, 1992-97.

Births (GMM Estimation)	Central Business Districts with Rent (I)	All Central Business Districts (II)
Office Rent	-0.044	..
	(0.196)	
Multi-county PMSA	..	-0.018
		(0.080)
Employment 1992	..	-0.052
		(0.116)
Wage	**-1.977**	**-1.660**
	(0.633)	(0.560)
SU scale	**1.384**	**1.348**
	(0.141)	(0.158)
SU size (median)	-0.009	-0.075
	(0.308)	(0.281)
Obs.	106	211
Sargan P-value	0.645	0.426

Coefficients in bold face are significant at 5 percent and in bold-italic are significant at 10 percent.

[3] The instruments are market potential in 1930, retail wage in 1930, local government expenditures in 1987, ocean dummy, share of housing units in large complexes, ratio of housing units to commercial establishments, advertising wages in 1977, and median size of agency in 1977.

Table 5: Separation and sorting using the sample of single unit (SU) movers, 1992-97.

Movers	**Conditional Logit**			**Conditional Logit + 2SLS[1]**		
	(I)	**(II)**	**(III)[2]**	**(IV)**	**(V)**	**(VI)[2]**
Interactive Terms				**Conditional Logit (Interactive terms)**		
Mover size×SU size	**0.136**	**0.180**	**0.150**	**0.181**	**0.181**	**0.201**
	(0.070)	(0.084)	(0.072)	(0.084)	(0.084)	(0.088)
Mover size×Wage	..	..	**0.309**	..	..	***0.310***
			(0.152)			(0.165)
Mover size×SU scale	..	..	0.020	..	..	0.026
			(0.029)			(0.074)
County Attributes				**2SLS[3] (county-specific fixed effects)**		
CBD dummy	**-1.119**	**-1.085**	-0.517	**-0.495**	***-0.338***	0.278
	(0.148)	(0.149)	(0.385)	(0.175)	(0.173)	(0.174)
Multi-county PMSA	-0.149	-0.161	0.204	***-0.269***	**-0.291**	0.206
	(0.168)	(0.168)	(0.432)	(0.146)	(0.147)	(0.150)
Employment 1992	-0.165	-0.199	-0.167	***-0.624***	-0.527	**-0.713**
	(0.148)	(0.150)	(0.148)	(0.334)	(0.347)	(0.336)
Wage	-0.030	0.076	***-1.239***	0.158	0.372	**-1.047**
	(0.277)	(0.282)	(0.650)	(0.333)	(0.341)	(0.343)
SU scale	**1.100**	**1.072**	**1.017**	**0.867**	**0.679**	**0.736**
	(0.131)	(0.133)	(0.177)	(0.290)	(0.277)	(0.291)
SU size (median)	-0.513	0.258	***-0.555***	**-0.854**	-0.123	**-0.906**
	(0.326)	(0.558)	(0.334)	(0.233)	(1.368)	(0.231)
SU size squared	..	***-0.403***	..	..	-0.350	..
		(0.217)			(0.576)	
Obs.	255,340	255,340	255,340	58,480 [166]	58,480 [166]	58,480 [166]
Pseudo R2 [Centered R2]	0.140	0.141	0.143	0.057 [0.379]	0.057 [0.380]	0.061 [0.405]
[Hansen-Sargan P-value]	..	..	..	0.342	0.100	0.337

Coefficients in bold face are significant at 5 percent and in bold-italic are significant at 10 percent.

[1] The values in squared brackets are from the instrumental variable estimation using the county-specific fixed effects.

[2] We include a complete set of interactive terms, but do not report the coefficients for the interaction of size with dummies and 1990 population.

[3] The instruments are market potential in 1930, retail wage in 1930, local government expenditures in 1987, ocean dummy, share of housing units in large complexes, to commercial establishments, advertising wages in 1977, and median size of agency in 1977.